

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2004, 21:46:38 ; Search time 6584 Seconds
(without alignments)
13039.491 Million cell updates/sec

Title: US-09-978-360A-32
Perfect score: 2356
Sequence: i atcttggcgccacagtgg.....aaccaaaaaaaaaaaaaa 2356

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2354.4	99.9	2421	3	HSM805188
2	2311	98.1	2355	3	CR593496
3	1059	44.9	1059	9	AY417612
4	1012.4	43.0	1088	1	AL581422
5	1003.4	42.6	1160	5	EX340456
6	966	41.0	1046	5	EX342893
7	962	40.8	1142	1	AL563768
8	954.8	40.5	1042	5	BQ059008
9	946.8	40.2	1111	5	EX375981
10	904.8	38.4	1083	5	EX361590
11	890	37.8	981	5	EX342894
12	883.6	37.5	901	5	BUS28500
13	874.4	37.1	1100	4	BM920550
14	858	36.4	1039	5	EX334124
15	856.2	36.3	1039	1	AL530521
16	851.2	36.1	983	5	EX334447
17	848.2	36.0	1332	3	AK044088
18	848.2	36.0	3683	3	AK083295
19	848.2	36.0	3691	3	AK076419
20	847.2	36.0	1019	4	BM544329
21	846.6	35.9	1352	3	AK002416
22	844.4	35.8	933	5	EX370758
23	840.2	35.7	1141	4	BM806153
24	839.4	35.6	946	5	BUI146835

```

25 837.2 35.5 1059 9 AY417614
26 837 35.5 913 5 BQ881855
27 835 35.4 845 5 BX364653
28 834.8 35.4 897 5 BU540823
29 825 35.0 885 5 BU542498
30 824.6 35.0 1087 5 EX402007
31 824.4 35.0 865 1 AL563662
32 815 34.6 985 5 BX354457
33 799.4 33.9 1067 4 BM922934
34 797.6 33.9 852 4 BM044182
35 793.8 33.7 980 1 AL559897
36 789.8 33.5 1028 5 EX402008
37 787.2 33.4 1036 4 BM561679
38 787 33.4 880 6 CA488482
39 786 33.4 935 5 BX334448
40 785 33.3 928 6 CA488412
41 780.2 33.1 816 4 BG680346
42 770.8 32.7 923 5 EX362069
43 765.8 32.5 873 4 BG681198
44 746.4 31.7 822 5 BX364654
45 745.6 31.6 868 4 BM048526

```

ALIGNMENTS

```

RESULT 1
LOCUS HSM805188 2421 bp mRNA linear HTC 03-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp762M0911 (from clone DKFZp762M0911).
ACCESSION AL833889
VERSION AL833889.1 GI:21739403
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2421)
AUTHORS Blum H., Bauersachs S., Mewes H.W., Weil B., Amid C., Osanger A.,
Fobo G., Han M. and Wiemann S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (03-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University, Munich/Germany)
within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp762M0911) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp762M0911
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

```

FEATURES

```

source
Location/Qualifiers
1..2421
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp762M0911"
/db_xref="taxon:9606"
/clone="DKFZp762M0911"
/tissue_type="melanoma (MeWo cell line)"
/clone_lib="762 (synonym: hmel2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/notes="hypothetical protein"
1..2421
/gene="DKFZp762M0911"
105..1163
/gene="DKFZp762M0911"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD38745.1"

```

gene

CDS

/db_xref="GI:21739404"
/db_xref="GOA:Q969V5"
/translation="MESGRPSLCOFILLTTSVVTAAALYSVVRQKARVSQELKGAKK
VHLDELKSLTSAPKGVYAVETLGAIVRSVETLNSQFVENCQVITQRLTLQEHKRV
WNRTHLWDCSKI IHORTNVPDLVPHEDVDVA VRVLKPLDSDVLDLETVEKPH
PSTQSDVDVGHVYISGRPKGI OETEMLVKVGATL TGVELVLDNNSVRLQPPKQMG
YILSSQFDSLLQRBSSVRLWKVLALVFPFCATILFFILRQYLRQELRLKMQ
EEFQHEBAQLLSRAKPDRESLSACVVCUSLFSKCVFLECGHVCSTECYRALPEPK
KCPICRQAITRVIPLYNS"

ORIGIN		Query Match	99.9%; Score 2354.4; DB 3; Length 2421;	
		Best Local Similarity	100.0%; Pred. No. 0;	
		Matches 2355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1	ATCCTTGGCGCCACAGTCGGCCACCGGGGCTCGCGCGCTCATGGAGCGGAGGGCGGC	60	
DB	64	ATCCTTGGCGCCACAGTCGGCCACCGGGGCTCGCGCGCTCATGGAGCGGAGGGCGGC	123	
QY	61	CCTCGCTGTCAGTTTATCTCTGGGCACCACTCTGTGTCACCGCCGCTGACT	120	
DB	124	CCTCGCTGTCAGTTTATCTCTGGGCACCACTCTGTGTCACCGCCGCTGACT	183	
QY	121	CCGTGTACCGCAGAGCCCGGGTCTCCAGAGCTCAAGGGAGCTAAAAAGTTCAAT	180	
DB	184	CCGTGTACCGCAGAGCCCGGGTCTCCAGAGCTCAAGGGAGCTAAAAAGTTCAAT	243	
QY	181	TGGGTGAAGATTTAAAGAGTATCTTTTCAGAGCTCCAGGAAATGCGTCTATGCTG	240	
DB	244	TGGGTGAAGATTTAAAGAGTATCTTTTCAGAGCTCCAGGAAATGCGTCTATGCTG	303	
QY	241	TTATAGAGGAGCTGTGCGGTCTGTAAAGAAACGCTTAAACGAGTTTGTGAAAAC	300	
DB	304	TTATAGAGGAGCTGTGCGGTCTGTAAAGAAACGCTTAAACGAGTTTGTGAAAAC	363	
QY	301	GCAAGGGGTAATTCAGGGGTGACCTTCAGGACCAAGATGTGTGGAATCAACA	360	
DB	364	GCAAGGGGTAATTCAGGGGTGACCTTCAGGACCAAGATGTGTGGAATCAACA	423	
QY	361	CCACCTTTGGAATGATTTCTCAAGATCATTCATCAGAGGACCAACACAGTGCCTTTG	420	
DB	424	CCACCTTTGGAATGATTTCTCAAGATCATTCATCAGAGGACCAACACAGTGCCTTTG	483	
QY	421	ACCTGGTGCCTCAGAGGATGCGGTGATGTGGCTGTGCGAGTGTGAAGCCCTGGAT	480	
DB	484	ACCTGGTGCCTCAGAGGATGCGGTGATGTGGCTGTGCGAGTGTGAAGCCCTGGAT	543	
QY	481	CAGTGGATCTGGTCTAGAGACTGTGTATGAGAAGTTCCACCCCTCGATTCACTCTTCA	540	
DB	544	CAGTGGATCTGGTCTAGAGACTGTGTATGAGAAGTTCCACCCCTCGATTCACTCTTCA	603	
QY	541	CCGATGTATCGGCCACTACATCAGCGGTGAGCGGCCCAAGGCATCCAGAGACCGAGG	600	
DB	604	CCGATGTATCGGCCACTACATCAGCGGTGAGCGGCCCAAGGCATCCAGAGACCGAGG	663	
QY	601	AGATGTGAAGTGGGGGCCACCTCAGAGGGTTGGCGAATGCTGCTGGACACAACT	660	
DB	664	AGATGTGAAGTGGGGGCCACCTCAGAGGGTTGGCGAATGCTGCTGGACACAACT	723	
QY	661	CTGTCCGCTGCAGCGCCCAACAAGGATGAGTACTATCTTAAGCAGCAGGACTTCG	720	
DB	724	CTGTCCGCTGCAGCGCCCAACAAGGATGAGTACTATCTTAAGCAGCAGGACTTCG	783	
QY	721	ACAGCTGTGTCAGAGCAGGAGTCAGGCTCTGGAAGAGTGTGGCGCTGGTTT	780	
DB	784	ACAGCTGTGTCAGAGCAGGAGTCAGGCTCTGGAAGAGTGTGGCGCTGGTTT	843	
QY	781	TTGGCTTTGCCACATGTGCCACCTCTTCTTCAATCTCCGGAAGAGTACTCTCAGCGGC	840	
DB	844	TTGGCTTTGCCACATGTGCCACCTCTTCTTCAATCTCCGGAAGAGTACTCTCAGCGGC	903	
QY	841	AGGAGCGCTGCGCTCAAGCAGATGAGGAGGATTTCCAGGAGCAGTGGCCAGCTGC	900	

DB	904	AGGAGCGCTGCGCTCAAGCAGATGACGAGGAGTTCCAGGAGCATGAGCCCGAGTGC	963	
QY	901	TGAGCCGAGCAACCGCTGAGGACAGGGAGAGTCTGAAGAGCGCTGTGTAGTGTCTGA	960	
DB	964	TGAGCCGAGCAACCGCTGAGGACAGGGAGAGTCTGAAGAGCGCTGTGTAGTGTCTGA	1023	
QY	961	GCAGCTTCAAGTCTCGCTCTTTCTGAGTGTGGGACGTTTGTCTGACCGAGTGTCT	1020	
DB	1024	GCAGCTTCAAGTCTCGCTCTTTCTGAGTGTGGGACGTTTGTCTGACCGAGTGTCT	1083	
QY	1021	ACCGCGCTTGCAGAGCCCAAGAGTGCCTATCTGCAGACAGCGCATCACCCGGGTGA	1080	
DB	1084	ACCGCGCTTGCAGAGCCCAAGAGTGCCTATCTGCAGACAGCGCATCACCCGGGTGA	1143	
QY	1081	TACCCCTGTACACAGACTAATAGTTTGAAGAGCCCAAGAGTGTGCTGGAAGCACCCCTG	1140	
DB	1144	TACCCCTGTACACAGACTAATAGTTTGAAGAGCCCAAGAGTGTGCTGGAAGCACCCCTG	1203	
QY	1141	CCCCCTTTTTCAGGAGATTTTATCTCGAGGCTTTTGGAGGACAGTGGTGGGGGTAGCTGT	1200	
DB	1204	CCCCCTTTTTCAGGAGATTTTATCTCGAGGCTTTTGGAGGACAGTGGTGGGGGTAGCTGT	1263	
QY	1201	CACCTCCAGGTATGATTGAGGAGGAAATTGGGTAGAACTCTCCAGACCCATGCTCCAA	1260	
DB	1264	CACCTCCAGGTATGATTGAGGAGGAAATTGGGTAGAACTCTCCAGACCCATGCTCCAA	1323	
QY	1261	TGGCAGGATGCTGCTTTCCACCTGAGAGGGACCCCTGTCCATGTGCAGGCTCATCAGA	1320	
DB	1324	TGGCAGGATGCTGCTTTCCACCTGAGAGGGACCCCTGTCCATGTGCAGGCTCATCAGA	1383	
QY	1321	GCCTCACCTCGGAGGATGCGGTGCGTCTCTCCAGAGGACAGATCAGTGCAGAGTGT	1380	
DB	1384	GCCTCACCTCGGAGGATGCGGTGCGTCTCTCCAGAGGACAGATCAGTGCAGAGTGT	1443	
QY	1381	ACTGAAATGCTCATCACTTAAGACCAAGCCAGTGTATCAGCAGCTTCTGTCTCTG	1440	
DB	1444	ACTGAAATGCTCATCACTTAAGACCAAGCCAGTGTATCAGCAGCTTCTGTCTCTG	1503	
QY	1441	TGCTCTCTGTTTTTTTCTGTTGTAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1500	
DB	1504	TGCTCTCTGTTTTTTTCTGTTGTAATCTGTTGTTGTTGTTGTTGTTGTTGTTG	1563	
QY	1501	GGGAGGAAGGCTGGGCCCCGAGTACAAACGAGTGTGGGTGCTGCTCCGAAAGACT	1560	
DB	1564	GGGAGGAAGGCTGGGCCCCGAGTACAAACGAGTGTGGGTGCTGCTCCGAAAGACT	1623	
QY	1561	CTGCGCAGCTTTTCTTTCTCTCATGCCCGGGAACAGTCTTCTTCTCAGAAATCT	1620	
DB	1624	CTGCGCAGCTTTTCTTTCTCTCATGCCCGGGAACAGTCTTCTTCTCAGAAATCT	1683	
QY	1621	CAGGCTGGGCGAGTCAACTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1680	
DB	1684	CAGGCTGGGCGAGTCAACTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1743	
QY	1681	CACGTGTGTGTAGAGGACAAAAGAGTGAATGAGTCAACATCCGCTTCTGCCAGATGT	1740	
DB	1744	CACGTGTGTGTAGAGGACAAAAGAGTGAATGAGTCAACATCCGCTTCTGCCAGATGT	1803	
QY	1741	CGGGGCCCCGGGCAACAGATTGAAAGAGATCATGTGAAGGCGAGTGGTTCAGGACGGCC	1800	
DB	1804	CGGGGCCCCGGGCAACAGATTGAAAGAGATCATGTGAAGGCGAGTGGTTCAGGACGGCC	1863	
QY	1801	TCTGTGTTTTCGCACTGGGCGCTGATTGAACTCTTCTTCTTCTTCTTCTTCTTCTTCT	1860	
DB	1864	TCTGTGTTTTCGCACTGGGCGCTGATTGAACTCTTCTTCTTCTTCTTCTTCTTCTTCT	1923	
QY	1861	CCTGTGTTTTCCTCTCTGAGAAATGAGGCGAGAGCCCTCGCTTCTGAGGACGCGAGT	1920	
DB	1924	CCTGTGTTTTCCTCTCTGAGAAATGAGGCGAGAGCCCTCGCTTCTGAGGACGCGAGT	1983	
QY	1921	TGGATGCCACTGGCTAGTGTCTGCTCTCAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCT	1980	
DB	1984	TGGATGCCACTGGCTAGTGTCTGCTCTCAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCT	2043	

```
QY 1981 AAGCAGCGGGTGGCAACCTGAGCATATGCCCTCTTGCGGCTCCCTCATCCAGCCCGTCG 2040
Db |||||||
QY 2044 AAGCAGCGGGTGGCAACCTGAGCATATGCCCTCTTGCGGCTCCCTCATCCAGCCCGTCG 2103
Db |||||||
QY 2041 CAGCTTTGACATCTTGCTGCTACTCATGCTGCTTCTCCCTTGCTTTACCCCTCCAGTATT 2100
Db |||||||
QY 2104 CAGCTTTGACATCTTGCTGCTACTCATGCTGCTTCTCCCTTGCTTTACCCCTCCAGTATT 2163
Db |||||||
QY 2101 ACCATTTGCCCTCACCTGCCCTTGGTGGAGCTTTTGTAGTGAAGACAGATGGGCTGTTT 2160
Db |||||||
QY 2164 ACCATTTGCCCTCACCTGCCCTTGGTGGAGCTTTTGTAGTGAAGACAGATGGGCTGTTT 2223
Db |||||||
QY 2161 TCCCCCACTCTGAGTAGTTGGAGGTACATACACAGCTCTCTTTTATTGCCCTTTCT 2220
Db |||||||
QY 2224 TCCCCCACTCTGAGTAGTTGGAGGTACATACACAGCTCTCTTTTATTGCCCTTTCT 2283
Db |||||||
QY 2221 GCCTCTGAATGTTCACTCTCGCTCTCTCTTTGTGCGAGGAGGAGGCTCCCTCAGGG 2280
Db |||||||
QY 2284 GCCTCTGAATGTTCACTCTCGCTCTCTCTTTGTGCGAGGAGGAGGCTCCCTCAGGG 2343
Db |||||||
QY 2281 GCCGACACTAGTAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db |||||||
QY 2344 GCCGACACTAGTAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2403
Db |||||||
QY 2341 AAAAAAAAAAAAAA 2356
Db |||||||
QY 2404 AAAAAAAAAAAAAA 2419
Db |||||||

RESULT 2
CR593496
LOCUS
DEFINITION
full-length cDNA clone CS0DL011YN14 of B cells (Ramos cell line)
ACCESSION
CR593496
VERSION
1 (bases 1 to 2355)
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 2355)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMWSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..2355
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL011YN14"
/tissue type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMWSPORT_6"
ORIGIN
Query Match 98.1%; Score 2311; DB 3; Length 2355;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1081 TACCCCTGTAAACAGCTAATAGTTTGAAGCGGCACAGCTTGACCTGGAGCACCCCTG 1140
Db 1125 TACCCCTGTAAACAGCTAATAGTTTGAAGCGGCACAGCTTGACCTGGAGCACCCCTG 1184
QY 1141 CCCCCTTTTTCAGGATTTTATCTCGAGGCTTTTGAGGAGCAGTGGTGGGGGTAGCTGT 1200
Db 1185 CCCCCTTTTTCAGGATTTTATCTCGAGGCTTTTGAGGAGCAGTGGTGGGGGTAGCTGT 1244
QY 1201 CACCTCCAGGTATGATTTGAGGAGGAAATTTGGGTAGAAAATCTTCAGAACCCATGCCCTCCAA 1260
Db 1245 CACCTCCAGGTATGATTTGAGGAGGAAATTTGGGTAGAAAATCTTCAGAACCCATGCCCTCCAA 1304
QY 1261 TGGCAGGATGCTGCTTTTCCACCTCAGAGGGAACCTGTCCATGTGCGACCTCATCAGA 1320
Db 1305 TGGCAGGATGCTGCTTTTCCACCTCAGAGGGAACCTGTCCATGTGCGACCTCATCAGA 1364
QY 1321 GCCTCACCTGGGAGGATGCGGTGGCTCTCTCCAGGAGCCAGATCAGTGGAGGTG 1380
Db 1365 GCCTCACCTGGGAGGATGCGGTGGCTCTCTCCAGGAGCCAGATCAGTGGAGGTG 1424
QY 1381 ACTGAAATGCCCTCATCACTTAAGCAACCAAGCCAGTGATCAGAGCTCTTCTGTTCTTG 1440
Db 1425 ACTGAAATGCCCTCATCACTTAAGCAACCAAGCCAGTGATCAGAGCTCTTCTGTTCTTG 1484
QY 1441 TGTCTTCTGTTTCTGTTGATCGTTGCTTGTGCTGTGACCTGGTGGAGGACTCAGAG 1500
Db 1485 TGTCTTCTGTTTCTGTTGATCGTTGCTTGTGCTGTGACCTGGTGGAGGACTCAGAG 1544
QY 1501 GGGAGGAAAGGCTGGGCCCCGAGTACAAACGATGCTTGGGTGCTGCTCCGAGAGACT 1560
Db 1545 GGGAGGAAAGGCTGGGCCCCGAGTACAAACGATGCTTGGGTGCTGCTCCGAGAGACT 1604
QY 1561 CTCGCGCAGCTTCTTCTTCTTCTTCTCATGCGCCCGGAAACAGTCTTCTTCTCAGATGT 1620
Db 1605 CTCGCGCAGCTTCTTCTTCTTCTTCTCATGCGCCCGGAAACAGTCTTCTTCTCAGATGT 1664
QY 1621 CAGCTGGGAGGCTCACTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
Db 1665 CAGCTGGGAGGCTCACTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1724
QY 1681 CAGCTGTGTAGAGGACAAAGAAAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1740
Db 1725 CAGCTGTGTAGAGGACAAAGAAAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1784
QY 1741 CGGGGCCCCGGGCAACAGATGAAGAGATCATGTGAAGGAGGAGTGGTCAAGGAGGAGG 1800
Db 1785 CGGGGCCCCGGGCAACAGATGAAGAGATCATGTGAAGGAGGAGTGGTCAAGGAGGAGG 1844
QY 1801 TCCTGGTTTCGCCACTGGCCCTGATTTGAATCTCTGCACTTGGGAGAGCTCGGGGTGGT 1860
Db 1845 TCCTGGTTTCGCCACTGGCCCTGATTTGAATCTCTGCACTTGGGAGAGCTCGGGGTGGT 1904
QY 1861 CCCTGGTTTTCCTCTCTGGAGATGAGCGCAGAGGCTCGCTCTCTTGAAGGAGCAGAGT 1920
Db 1905 CCCTGGTTTTCCTCTCTGGAGATGAGCGCAGAGGCTCGCTCTCTTGAAGGAGCAGAGT 1964
QY 1921 TGGATGCCACTGGCTAGTCTGCGCTCAAGCTTCTTGAAGGAGTCAAGGAGTCAAGGAG 1980
Db 1965 TGGATGCCACTGGCTAGTCTGCGCTCAAGCTTCTTGAAGGAGTCAAGGAGTCAAGGAG 2024
QY 1981 AAGCAGCGGTGGCACCCTGAGCATATGCCCTCTTGGGGCTCCCTCATCCAGCCCGTCG 2040
Db 2025 AAGCAGCGGTGGCACCCTGAGCATATGCCCTCTTGGGGCTCCCTCATCCAGCCCGTCG 2084
QY 2041 CAGCTTTGACATCTTGTGTACTCATGTGCTTCTTCTTGTGTATCCCTCCAGTAT 2100
Db 2085 CAGCTTTGACATCTTGTGTACTCATGTGCTTCTTCTTGTGTATCCCTCCAGTAT 2144
QY 2101 ACCATTTGCCCTCACCTGCCCTTGGTCAAGCTTTTGTAGTCAAGCAGATGGGCTGTT 2160
Db 2145 ACCATTTGCCCTCACCTGCCCTTGGTCAAGCTTTTGTAGTCAAGCAGATGGGCTGTT 2204
QY 2161 TCCCCCACTCTGAGTGTGGAGGTACATACACAGCTCTTTTTTTTATTGCCCCCTTTCT 2220
```

```
Db 2205 TCCCCCACTCTGAGTAGTTGGAGGTACATACAGCTCTTTTTTTATTGCCCCCTTTCT 2264
QY 2221 GCTCTGTAATGTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280
Db 2265 GCTCTGTAATGTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2324
QY 2281 GCGCACACTAGTATGATGACAGTGTCCAGTGT 2311
Db 2325 GCGCACACTAGTATGATGACAGTGTCCAGTGT 2355

RESULT 3
AY417612 1059 bp DNA linear GSS 12-DEC-2003
LOCUS Homo sapiens HCM6282 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY417612
VERSION AY417612.1 GI:39773572
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1059)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1059)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1059
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1059
/locus_tag="HCM6282"
ORIGIN
Query Match 44.9%; Score 1059; DB 9; Length 1059;
Best Local Similarity 100.0%; Pred. No. 4,3e-269;
Matches 1059; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 ATGAGAGAGCGAGGCGCGCTCGCTGTCAGTTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 101
Db 1 ATGAGAGAGCGAGGCGCGCTCGCTGTCAGTTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 60
QY 102 GTCAACCGCGCCCTGTACTCCGTGTACCGGAGAGGCCCGGGTCTCCCAAGAGCTCAAG 161
Db 61 GTCAACCGCGCCCTGTACTCCGTGTACCGGAGAGGCCCGGGTCTCCCAAGAGCTCAAG 120
QY 162 GGAGCTAAAAAGTTCAATTTGGGTGAAGATTAAAGAGATTCTTTCAGAGCTCCAGGA 221
Db 121 GGAGCTAAAAAGTTCAATTTGGGTGAAGATTAAAGAGATTCTTTCAGAGCTCCAGGA 180
QY 222 AAATGCGTGCCTTATGCTGTATAGAAGGAGCTGTGGGCTGTGTTAAAGAAACGCTTAAC 281
Db 181 AAATGCGTGCCTTATGCTGTATAGAAGGAGCTGTGGGCTGTGTTAAAGAAACGCTTAAC 240
QY 282 AGCCAGTTTGTGGAAAACACTGCAAGGGGGTAATTTCAGCGGCTGACACCTTCAGGAGCACAG 341
```

```

Db 241 AGCCAGTTTGTGAAAACCTGCAAGGGGTAAATTCAGCGCTCACATTCAGAGACACAG 300
QY 342 ATGGTGTGAATCGAACACCCACCTTTTGGATGATGCTCAAGATCATTCATCAGAGG 401
Db 301 ATGGTGTGAATCGAACACCCACCTTTTGGATGATGCTCAAGATCATTCATCAGAGG 360
QY 402 ACCAACACAGTGCCTTTTACCTGTGTGCCCCACGAGGATGGCGTGTGATGCGCA 461
Db 361 ACCAACACAGTGCCTTTTACCTGTGTGCCCCACGAGGATGGCGTGTGATGCGCA 420
QY 462 GTGCTGAAGCCCTCGACTCAGTGGATCTGGGTCTAGAGACTGTGTATGAGAAGTTCCAC 521
Db 421 GTGCTGAAGCCCTCGACTCAGTGGATCTGGGTCTAGAGACTGTGTATGAGAAGTTCCAC 480
QY 522 CCCTCGATTCAGTCTTACCCATGTATCGCGCCATACATCAGCGGTGAGCGGCCCAA 581
Db 481 CCCTCGATTCAGTCTTACCCATGTATCGCGCCATACATCAGCGGTGAGCGGCCCAA 540
QY 582 GGCATCCAGAGACCGAGGAGATGCTGAAGGTGGGGCCACCTCACAGGGGTGGCGAA 641
Db 541 GGCATCCAGAGACCGAGGAGATGCTGAAGGTGGGGCCACCTCACAGGGGTGGCGAA 600
QY 642 TTGGTCTGGACCAACAATCTGTCTCCGCTGCGAGCCGCCCAACAAGGCATGCACTAT 701
Db 601 TTGGTCTGGACCAACAATCTGTCTCCGCTGCGAGCCGCCCAACAAGGCATGCACTAT 660
QY 702 CTAAGCAGCAGGACTTCGACAGCTGTCTGACAGGCGAGGATCGAGCTCAGGCTTGG 761
Db 661 CTAAGCAGCAGGACTTCGACAGCTGTCTGACAGGCGAGGATCGAGCTCAGGCTTGG 720
QY 762 AAGTGTCTGGCGCTGGTTTGGCTTGGCCATGTGCCACCTCTCTTCATCTCCGG 821
Db 721 AAGTGTCTGGCGCTGGTTTGGCTTGGCCATGTGCCACCTCTCTTCATCTCCGG 780
QY 822 AAGCAGTATCTGACGCGCAGGAGCGCTTGGCCCTCAAGCAGATCAGGAGGATTCAG 881
Db 781 AAGCAGTATCTGACGCGCAGGAGCGCTTGGCCCTCAAGCAGATCAGGAGGATTCAG 840
QY 882 GAGCATGAGGCCAGCTGCTGAGCGCAGCAGCAAGCTTGAAGCAGGAGGATCTGAAGAGC 941
Db 841 GAGCATGAGGCCAGCTGCTGAGCGCAGCAGCAAGCTTGAAGCAGGAGGATCTGAAGAGC 900
QY 942 GCCTGTGTAGTGTCTGAGCAGCTTCAAGTCTCGCTTCTTCTGAGTGTGGGCACGTT 1001
Db 901 GCCTGTGTAGTGTCTGAGCAGCTTCAAGTCTCGCTTCTTCTGAGTGTGGGCACGTT 960
QY 1002 TGTTCCTGACCGAGTGTCTACCGCGCTTGGCCAGAGCCCAAGAGTGCCTATCTGCAGA 1061
Db 961 TGTTCCTGACCGAGTGTCTACCGCGCTTGGCCAGAGCCCAAGAGTGCCTATCTGCAGA 1020
QY 1062 CAGGCGATCACCAGGATGATACCCCTGTACACAGCTAA 1100
Db 1021 CAGGCGATCACCAGGATGATACCCCTGTACACAGCTAA 1059

```

RESULT 4

```

AL581422/c 1088 bp mRNA linear EST 07-APR-2004
LOCUS AL581422 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0DG003YC07 3-PRIME, mRNA sequence.
ACCESSION AL581422
VERSION AL581422.3 GI:46259998
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1088)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:31319662.

```

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

This sequence belongs to sequence cluster 670.r

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?b=CS0DG003AB04NFP1&c=670.r>.

FEATURES

source

1..1088

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DG003YC07"

/tissue_type="B CELLS (RAMOS CELL LINE)"

/cell_line="RAMOS CELL LINE"

/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

```

Query Match 43.0%; Score 1012.4; DB 1; Length 1088;
Best Local Similarity 93.9%; Pred. No. le-256;
Matches 1024; Conservative 42; Mismatches 21; Indels 3; Gaps 2;

QY 1237 AACTCTCAGACCCATGCTTCAATGGCAGGATGCTGCTTCCACCTGAGGGGACC 1296
Db 1088 AAATTTCCMAAMCCYTCYCMATKSGSAGATBT--CYTTTCCMCYTGAGGGGCC 1031
QY 1297 CTGTCATGTGCAGCTCATCAGAGCCTCACCTGGGAGGATGCGTGCGTCTCTCC 1356
Db 1030 YTKYCHWTGTSCACCTTCATCAGAGCCTCACCTGGGAGGATGCGTGCGTCTCTCC 971
QY 1357 AGGAGCCAGATCAGTGGAGTGTGACTGAAATGCTCATCTTAAGCACCACCAAGCCAG 1416
Db 970 AGGAGCCAGATCAGTGGAGTGTGACTGAAATGCTCATCTTAAGCACCACCAAGCCAG 911
QY 1417 TGATCAGAGCTCTTCTGCTGCTGCTTCTGTTTTTTCTGTTGAATGCTGCTGCT 1476
Db 910 TGATCAGAGCTCTTCTGCTGCTGCTTCTGTTTTTTCTGTTGAATGCTGCTGCT 851
QY 1477 GTGGACTTGTGGAGGACTCAGAGGGAGGAGGAGCTGGGCCCCGAGTACACGATGCC 1536
Db 850 GTGGACTTGTGGAGGACTCAGAGGGAGGAGGAGCTGGGCCCCGAGTACACGATGCC 791
QY 1537 TTGGGTGCTGCTCCGAAGAGACTCTGCCGAGACTTTTCTTCTTCTCATGCCCCGG 1596
Db 790 TTGGGTGCTGCTCCGAAGAGACTCTGCCGAGACTTTTCTTCTTCTCATGCCCCGG 731
QY 1597 GAAACAGCTTTCTTCAGATTGTGAGCTGGCAGGTCAACTTGTGTCTTCTTCTCCCTC 1656
Db 730 GAAACAGCTTTCTTCAGATTGTGAGCTGGCAGGTCAACTTGTGTCTTCTTCTCCCTC 671
QY 1657 ACCTGCTGCTCCTTAACGCTGCGACGTGTGTAGAGGACAAAGAAAGTGAAGTCAG 1716
Db 670 ACCTGCTGCTCCTTAACGCTGCGACGTGTGTAGAGGACAAAGAAAGTGAAGTCAG 611
QY 1717 CACATCCGCTTCTGCCCCAGATGCTGGGGCCCCGGGCAACAGATTGAAGAGAGATCATGT 1776
Db 610 CACATCCGCTTCTGCCCCAGATGCTGGGGCCCCGGGCAACAGATTGAAGAGAGATCATGT 551
QY 1777 GAAAGGAGATTTGGTCAGCAGCCCTCTCTGTTTTCGCCACCTGGCCCTGATTTGAATCTCTG 1836
Db 550 GAAAGGAGATTTGGTCAGCAGCCCTCTCTGTTTTCGCCACCTGGCCCTGATTTGAATCTCTG 491
QY 1837 CCATTTGGGAGAGCTCGGGGTGGTCCCTTGTTTTCCTCTCTGGAGATGAGCGCGAGAGG 1896

```

[illegible]

RESULT 5
BX340456/c
LOCUS
DEFINITION
BX340456 1160 bp mRNA linear EST 08-APR-2004
BX340456 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CS0DI087YH04 3-PRIME, mRNA sequence.
ACCESSION
VERSION BX340456.2 GI:46282225
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1160)
REFERENCE
Li,W.B., Gruber,C., Jesse,J., and Polayes,D.
Full-length CDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30345890.

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 670.r
 For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CS0DI087DD02NP1&c=670.r>.
 Location/Qualifiers
 1. 1160
 source

ORIGIN		/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."									
	Query Match	42.6%	Score 1003.4	DB 5	Length 1160						
	Best Local Similarity	95.0%	Pred. No. 2.5e-254								
	Matches 1073	Conservative 12	Mismatches 36	Indels 8	Gaps 5						
Qy	1183	AGTGGTGGGGTAGCTGTCTACCTCCAGGTATGATTGAGGGAGAAATGGGTAGAAACTCT	1242								
Db	1121	RGGGCAGTGGGGTGTGTTTACCCAGTGTGATTGAGGARG--AATGGGGAAAAACTCT	1064								
Qy	1243	CCAGACCATGCCTCAAATGGCAGGATGCTGCCTTTCCACCTGAGAGGGAGCCCTGTGCC	1302								
Db	1063	CCARACCATCCTCCAAATGCAGA--GGYTCCTTTCCACCTTGAG--GGGACCTGTGCC	1007								
Qy	1303	ATGTGACGCTCATCAGAGCCTCACCTGGGAGGATGCCTGTGGCGTCTCTCCACGAGGC	1362								
Db	1006	ATGTGACGCTCATCAGAGCCTCACCTGGGAGGATGCCTGTGGCGTCTCTCCACGAGGC	948								
Qy	1363	CAGATCAGTCGAGTGTGACTGAAAATGCCTCATCACTTAAGCACCAAGACCAAGTATCA	1422								
Db	947	CAGATCAGTCGAGTGTGACTGAAAATGCCTCATCACTTAAGCACCAAGACCAAGTATCA	898								
Qy	1423	GCAGCTCTTCGTCTGCTGCTCTCTGTTTTTTCTTGGTGAATGTTGCTGTGTGGAC	1482								
Db	887	GCAGCTCTTCGTCTGCTGCTCTCTGTTTTTTCTGTTGTCGTCGCTGTGTGTGGAC	828								
Qy	1483	TTGGTGAGGACTCAGAGGGGAGAAAGGCTGGGCCCCCGAGTCAACGGATGCTTTGGGT	1542								
Db	827	TTGGTGAGGACTCAGAGGGGAGAAAGGCTGGGCCCCCGAGGACCAACGGATGCTTTGGGT	768								
Qy	1543	GCTGCTCCGAGAGACTCTGCGCGAGCTTTTCTTCTTTTCTCATGCCCCCGGAAACA	1602								
Db	767	GCTGCTCCGAGAGACTCTGCGCGAGCTTTTCTTCTTTTCTCATGCCCCCGGAAACA	708								
Qy	1603	GTCTTTCTTCAGAAATGTCAGGCTGGGCGAGTCAACTGTGTTCCTTTCCCTCACCTGC	1662								
Db	707	GTCTTTCTTCAGAAATGTCAGGCTGGGCGAGTCAACTGTGTTCCTTTCCCTCACCGGC	648								
Qy	1663	TTGCCTCCTTAAACGCTTCGCTGTGTGTAGAGGACAAAAGTGAAGTCAAGCACATC	1722								
Db	647	TTGCCTCCTTAAACGCTTCGCTGTGTGTAGAGGACAAAAGTGAAGTCAAGCACATC	588								
Qy	1723	CGCTTCGCCAGATGTGCGGGCCCCCGGCGAACAGATTGAAGAGATCATGTGAAGGG	1782								
Db	587	CGCTTCGCCAGATGTGCGGGCCCCCGGCGAACAGATTGAAGAGATCATGTGAAGGG	528								
Qy	1783	CAGTTGCTCAGGACGGCTCCTGGTTTCGCCACTGGCCCTGATTGGAACCTCTGCCACTT	1842								
Db	527	CAGTTGCTCAGGACGGCTCCTGGTTTCGCCACTGGCCCTGATTGGAACCTCTGCCACTT	468								
Qy	1843	GGGAGAGCTCGGGGTGTCTGTTTTCCTCTCTGGAGAAATGAGCGCGAGAGCCCTCGC	1902								
Db	467	GGGAGAGCTCGGGGTGTCTGTTTTCCTCTCTGGAGAAATGAGCGCGAGAGCCCTCGC	408								
Qy	1903	CTCTCTGAAGACGCGAGTGTGGATGCCACTGGCCTAGTGTCTCTGGCCTCAGAGCTTCTTG	1962								
Db	407	CTCTCTGAAGACGCGAGTGTGGATGCCACTGGCCTAGTGTCTCTGGCCTCAGAGCTTCTTG	348								
Qy	1963	CAGGCTGTCAAGGAAAAGCAGCGGCTGGCACCTGAGCATATGCCCTCTTGGGGCT	2022								
Db	347	CAGGCTGTCAAGGAAAAGCAGCGGCTGGCACCTGAGCATATGCCCTCTTGGGGCT	288								
Qy	2023	CCCTCATCCAGCCGTCGCACTTTGACATCTTGGTGTACTGTCTGCTTCTCTTTGTG	2082								
Db	287	CCCTCATCCAGCCGTCGCACTTTGACATCTTGGTGTACTGTCTGCTTCTCTTTGTG	228								
Qy	2083	TTACCCCTCCAGTATTACATTGTCCTCTCACTGCCCTTGGTGAGCCCTTTAGTGCA	2142								
Db	227	TKVCCCTCCAGTATTACATTGTCCTCTCACTGCCCTTGGTGAGCCCTTTAGTGCA	168								

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1142)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31287753.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
670.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DD007AD08NP1&c=670.r.

FEATURES

source

Location/Qualifiers

1..1142
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD007V315"
/tissue="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 40.8%; Score 962; DB 1; Length 1142;
Best Local Similarity 89.8%; Pred. No. 2.3e-243;
Matches 1025; Conservative 40; Mismatches 72; Indels 5; Gaps 4;

QY 1184 GTGGTGGGGTACTGTACCTCAGTATGATTGAGGAGGAATTGGGTAGAACTCTC 1243
DB 1137 GCGGGTAKGGGGGKTKTTTCCGCCCAATTTTATKARGCGAGAAWTTGGGAAAAATY 1078
QY 1244 CAGACCCATCCCTCCAAATGCGAGATGCTGCCCTCCACCTGAGAGGACCTGTCCA 1303
DB 1077 TCMAAACCATCTCCAAATGCGAGAGTCTCCTTTTCCACCTTARRGGAACTKGCCT 1020
QY 1304 TGTGCGACCTCATCAGAGCCTCACCTGGGAGGATGCGGTGGGTCTCTCCAGGAGCC 1363
DB 1019 TG-GGACCTCTWYAAACCTTAACCTGGGAGGA-XSCGTGKGTCTCTCTCCAGGAGCC 962
QY 1364 AGATCAGTGGAGTGAGTGAATGCTCATCATCTTAACCAACCAAGCGATGATCAG 1423
DB 961 AGATCAGTGGAGTGAATGCTCATCATCTTAACCAACCAAGCGATGATCAG 903
QY 1424 CAGCTCTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1483
DB 902 CAGCTCTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 843
QY 1484 TGTGAGGAGTCTCAGAGGGAGGAAAGGCTGGGCGCCGAGTACCAACCGATGCTTGGGTG 1543
DB 842 TGTGAGGAGTCTCAGAGGGAGGAAAGGCTGGGCGCCGAGTACCAACCGATGCTTGGGTG 783
QY 1544 CTGCTCCGAGAGTCTGCGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1603
DB 782 CTGCTCCGAGAGTCTGCGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 723
QY 1604 TCTTCTTTCAGAAATGTGAGGCTGGGAGGTCACTTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1663
DB 722 TCTTCTTTCAGAAATGTGAGGCTGGGAGGTCACTTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 663
QY 1664 TGCCTCTTACCGCTGCACTGT 1723
DB 662 TGCCTCTTACCGCTGCACTGT 603

FEATURES
source

```

Location/Qualifiers
1. .1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5814982"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

Query Match      40.5%; Score 954.8; DB 5; Length 1042;
Best Local Similarity 97.3%; Pred. No. 1.8e-241;
Matches 990; Conservative 0; Mismatches 25; Indels 2; Gaps 2;

QY 984 CTGGAGTGGGACGCTTTGTTCTGCACCGAGTGTACCGCGCCTTGCCAGAGCCCAAG 1043
Db 26 CTGGAGTGGGACGCTTTGTTCTGCACCGAGTGTACCGCGCCTTGCCAGAGCCCAAG 85

QY 1044 AAGTGCCTTATCTGCAGACAGGCGATCACCCGGGTGATACCCCTGTACACAGCTAATAG 1103
Db 86 AAGTGCCTTATCTGCAGACAGGCGATCACCCGGGTGATACCCCTGTACACAGCTAATAG 145

QY 1104 TTGTGAAGCGGCACACTTGACCTGGAAGCACCCTGCCCTTTTCAGGATTTTATC 1163
Db 146 TTGTGAAGCGGCACACTTGACCTGGAAGCACCCTGCCCTTTTCAGGATTTTATC 205

QY 1164 TCGAGGCTTTGGAGGAGCAGTGGTGGGGTAGCTGTCACTCCAGGTATGATTGAGGA 1223
Db 206 TCGAGGCTTTGGAGGAGCAGTGGTGGGGTAGCTGTCACTCCAGGTATGATTGAGGA 265

QY 1224 GGAATGGGTAGAACTCTCGACCCATGCTTCAATGGGAGGAGTGCCTTTCCAC 1283
Db 266 GGAATGGGTAGAACTCTCGACCCATGCTTCAATGGGAGGAGTGCCTTTCCAC 325

QY 1284 CTGAGAGGGGACCTCTGCATGTGAGCTCATCAGAGCTCACCTGGGAGGATCCCT 1343
Db 326 CTGAGAGGGGACCTCTGCATGTGAGCTCATCAGAGCTCACCTGGGAGGATCCCT 385

QY 1344 GCGCTCTCTCCAGAGCCAGATCAGTGCAGTGTGACTGAAATGCTCATCACTTAA 1403
Db 386 GCGCTCTCTCCAGAGCCAGATCAGTGCAGTGTGACTGAAATGCTCATCACTTAA 445

QY 1404 GCACCAAGCCAGTATCAGAGCTCTTCTGTTCTGTGCTCTCTGTTTTTCTGGTGA 1463
Db 446 GCACCAAGCCAGTATCAGAGCTCTTCTGTTCTGTGCTCTCTGTTTTTCTGGTGA 505

QY 1464 ATCGTTGCTTGTGAGCTTGGTGGAGCTCAGAGGGAGGAGGCTGGGCCCCGAG 1523
Db 506 ATCGTTGCTTGTGAGCTTGGTGGAGCTCAGAGGGAGGAGGCTGGGCCCCGAG 565

QY 1524 TACAACGAGTGCCTTGGGTGCTGCTCCGAAAGAGACTCTGCGCGAGCTTTCTCTTTT 1583
Db 566 TACAACGAGTGCCTTGGGTGCTGCTCCGAAAGAGACTCTGCGCGAGCTTTCTCTTTT 625

QY 1584 CTTATGCCCCGGGAAACAGTCTTTCTTCAGAAATTTGTAGGCTGGGAGGTCAACTTGTG 1643
Db 626 CTTATGCCCCGGGAAACAGTCTTTCTTCAGAAATTTGTAGGCTGGGAGGTCAACTTGTG 685

QY 1644 TTCTTTTCCCTTCACCTGCTTCCCTCCCTTAACGCTGCACGCTGTGTGTAGAGGACAAAG 1703
Db 686 TTCTTTTCCCTTCACCTGCTTCCCTCCCTTAACGCTGCACGCTGTGTGTAGAGGACAAAG 745

QY 1704 AAAGTGAAGTCAGCACATCCCGTCTCTGCCAGATGTTGGGGCCCCCGGCAACAGATTGA 1763

```

Db

```

746 AAAGTGAAGTCAGCACATCCGCTTCTGCCAGATGTTGCGGGGCCCGGCAACAGATTGA 805
QY 1764 AGAGAGATCATGTGAAGGGGAGTGGTTCAGGAGGCTCTGTTTCGCCACTGGCCCTG 1823
Db 806 AGAGAGATCATGTGAAGGGGAGTGGTTCAGGAGGCTCTGTTTCGCCACTGGCCCTG 865
QY 1824 ATTGTAACCTCTCCACTTGGGAGAGCTCGGGTGTGTCCTGTTTCCCTCTCTGGAGAA 1883
Db 866 ATTGTAACCTCTCCACTTGGGAGAGCTCGGGTGTGTCCTGTTTCCCTCTCTGGAGAA 925
QY 1884 TGAAGGCGCAGAGGCTCGCTCTCTGAAGGAGCAGTGTGATGCCACTGGCCTAGT-GTC 1942
Db 926 TGAAGGCGCAGAGGCTCGCTCTCTGAAGGAGCAGTGTGATGCCACTGGCCTAGTGGTC 985
QY 1943 CTGGCCTCAC-ACGTTCTTGTCAAGCTGTCAAGGAAAGCAGCGGCTGGCACC 1998
Db 986 CTGGCCTCCCACTTCTTGGAGGTTGTCCCAAGAAACAGCCGCTGGGCC 1042

RESULT 9
BX375981/c
LOCUS
DEFINITION
BX375981 Homo sapiens NEUROBLASTOMA 3-PRIME, mRNA sequence.
ACCESSION
BX375981
VERSION
BX375981.2 GI:46574140
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1111)
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
On May 8, 2003 this sequence version replaced gi:30450430.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
670.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?S=CS0DC020DA04NP1&c=670.r.
Location/Qualifiers
1. .1111
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC020YB08"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
Query Match      40.2%; Score 946.8; DB 5; Length 1111;
Best Local Similarity 97.8%; Pred. No. 2.5e-239;
Matches 981; Conservative 8; Mismatches 11; Indels 3; Gaps 3;

QY 1288 GAGGGGACCTGTCTCATGTGAGGCTCATCAGAGCTCACCTGGGAGGATGCCGTGGCG 1347
Db 1001 GRGAGGGACCTKTCATGTGM-GCCTCATCAGAGCTTMACCTGGGAGGGTGGCGT-GCG 944
QY 1348 TCTCTCCAGAGGCCAGATCAGTGCAGTGTGACTGAAATGCTCATCACTTAAGCAC 1407

```



```
QY 541 CCGATGTCATCGGCCACTACATCAGCGGTGAGCGGCCCAAGGCATCCCAAGAGACCGAGG 600
Db 617 CCGATGTCATCGGCCACTACATCAGCGGTGAGCGGCCCAAGGCATCCCAAGAGACCGAGG 676
QY 601 AGATGCTGAAGGTGGGGGCCACCTCACA-GGGTTTGGCGAACTGCTCTGGACAAACAAC 659
Db 677 AGATGCTGAAGGTGGGGGCCACCTCACAAGGGGTTGGCGAACTGCTCTGGACAAACAAC 736
QY 660 TCTGTCGGCTCGAGCGGCCCAACAAGCATGCACTACTATCTAAGCAGCCAGGACTTC 719
Db 737 TCTGTCGGCTCGAGCGGCCCAACAAGCATGCACTACTATCTAAGCAGCCAGGACTTC 796
QY 720 GACAGCTCTGTCAGAGGAGGAGTCCAGCTCAGGCTCTGGAAGGTGCTGGCGCTGGTT 779
Db 797 GACAGCTCTGTCAGAGGAGGAGTCCAGCTCAGGCTCTGGAAGGTGCTGGCGCTGGTT 856
QY 780 TTTGGCTTTGCCACATGTGCCACCTCTCTTCATTCTCCGGAAGCAGTATCTGCAGCGG 839
Db 857 TTTGGCTTTGCCACATGTGTGCCACCTCTCTTCATTCTCCGGAAGCAGTATCTGCAGCGG 916
QY 840 CAGGAGCGCTCGCTCAGCAGATGTCAGGAGGAGTTCAGGAGCATGAGGCCAGCTG 899
Db 917 CAGGAGCGCTCGCTCAGCAGATGTCAGGAGGAGTTCAGGAGCATGAGGCCAGCTG 975
QY 900 CTGAGCGGAGCAAGCTCGAGGAGGAGTCTGAAGAGCGCTGTGTAGTGTCTG 959
Db 976 CTGAGCGGAGCAAGCTCGAGGAGGAGTCTGAAGAGCGCTGTGTAGTGTCTG 1035
QY 960 AGC 962
Db 1036 AGC 1038

RESULT 11
BX342894
LOCUS
DEFINITION
BX342894 Homo sapiens B CELLS (RAMOS CELL LINE) EST 07-APR-2004
Homo sapiens cDNA clone CS0DL007YCL16 5-PRIME, mRNA sequence.
ACCESSION
BX342894
VERSION
BX342894.2 GI:46275083
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 981)
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30340129.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
670.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DL007BB08QPl&c=670.r.
Location/Qualifiers
1..981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL007YCL16"
/cell_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
```

FEATURES

Source

primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

```
Query Match 37.8%; Score 890; DB 5; Length 981;
Best Local Similarity 99.6%; Pred. No. 2.8e-224;
Matches 923; Conservative 0; Mismatches 1; Indels 3; Gaps 3;
QY 1 ATCTTTGGCGCCACAGTCGGGCCACCGGGGCTCGCCCGCTCATGGAGAGCGAGGCGGCG 60
Db 57 ATCTTTGGCGCCACAGTCGGGCCACCGGGGCTCGCCCGCTCATGGAGAGCGAGGCGGCG 116
QY 61 CTCTGCTGTGCGAGTTTCCTCTCTGGGCAACCACTCTCTGTGTCTACCGCGCGCTTACT 120
Db 117 CTCTGCTGTGCGAGTTTCCTCTCTGGGCAACCACTCTCTGTGTCTACCGCGCGCTTACT 176
QY 121 CCGTGTACCGGAGAGGCGCGGGTCTCCCAAGAGCTCAAGGAGCTAAAGAGTTCAAT 180
Db 177 CCGTGTACCGGAGAGGCGCGGGTCTCCCAAGAGCTCAAGGAGCTAAAGAGTTCAAT 236
QY 181 TGGGTGAAGATTTAAAGAGTATTTCTTTAGAAAGCTCCAGGAAATCGCTTATGCTG 240
Db 237 TGGGTGAAGATTTAAAGAGTATTTCTTTAGAAAGCTCCAGGAAATCGCTTATGCTG 296
QY 241 TTATAGAGGAGCTGTGCGGTCTGTTAAAGAAACGCTTAAACAGCCAGTTTGTGAAACT 300
Db 297 TTATAGAGGAGCTGTGCGGTCTGTTAAAGAAACGCTTAAACAGCCAGTTTGTGAAACT 356
QY 301 GCAAGGGGTAAATTCAGGCGCTGACACTTCAGGAGCAAGATGTTGTAATCGAACCA 360
Db 357 GCAAGGGGTAAATTCAGGCGCTGACACTTCAGGAGCAAGATGTTGTAATCGAACCA 416
QY 361 CCCACCTTTGGAATGATTTGCTCAAGAGTCAATTCATCAGAGGACCAACACAGTGCCTTCA 420
Db 417 CCCACCTTTGGAATGATTTGCTCAAGAGTCAATTCATCAGAGGACCAACACAGTGCCTTCA 476
QY 421 ACTGTGTCGCCACAGAGATGCGGTGTGTGGTGTGTGAGTGTGAGAGCCCTTGGACT 480
Db 477 ACTGTGTCGCCACAGAGATGCGGTGTGTGGTGTGTGAGTGTGAGAGCCCTTGGACT 536
QY 481 CAGTGTGTCGGTCTAGAGACTGTGTAGAAAGTTTCCACCCCTCGATTTCAGTCTCTTCA 540
Db 537 CAGTGTGTCGGTCTAGAGACTGTGTAGAAAGTTTCCACCCCTCGATTTCAGTCTCTTCA 596
QY 541 CCGATGTCATCGGCCACTACATCAGCGGTGAGCGGCCCAAGGCATCCCAAGAGACCGAGG 600
Db 597 CCGATGTCATCGGCCACTACATCAGCGGTGAGCGGCCCAAGGCATCCCAAGAGACCGAGG 656
QY 601 AGATGCTGAAGGTGGGGGCCACCTCACAAGGGTGGCGAACTGCTCTGGACAAACAAC 660
Db 657 AGATGCTGAAGGTGGGGGCCACCTCACAAGGGTGGCGAACTGCTCTGGACAAACAAC 716
QY 661 CTGTCGGCTCGAGCGGCCCAACAAGCATGCACTACTATCTAAGCAGCCAGGACTTCG 720
Db 717 CTGTCGGCTCGAGCGGCCCAACAAGCATGCACTACTATCTAAGCAGCCAGGACTTCG 776
QY 721 ACAGCTCTGTCAGAGGAGGAGTCCAGCTCAGGCTCTGGAAGGTGCTGGCGCTGGTT 780
Db 777 ACAGCTCTGTCAGAGGAGGAGTCCAGCTCAGGCTCTGGAAGGTGCTGGCGCTGGTT 836
QY 781 TTGGCTTTGCCACATGTGCGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 839
Db 837 TTGGCTTTG-CACATGTGCGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 895
QY 840 CAGGAGCGCTCGCGCTCAAGCAGATGCAAGGAGGTTCCAGGAGCATGAGGCCAGCTG 899
Db 896 CAGGAGCGCTCGCGCTCAAGCAGATGCAAGGAGGTTCCAGGAGCATGAGGCCAGCTG 954
QY 900 CTGAGCGGAGCAAGCTGTGAGGACAGG 926
Db 955 CTGAGCGGAGCAAGCTGTGAGGACAGG 981
```

```
RESULT 12
BU528500
LOCUS
DEFINITION
AGENCOURT 10182994 NIH_MGC_101 Homo sapiens cDNA clone
IMAGE:6538449 5', mRNA sequence.
ACCESSION
BU528500
VERSION
BU528500.1 GI:22838941
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 901)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2703 row: a column: 09
High quality sequence stop: 660.
Location/Qualifiers
1..901
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6538449"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/notes="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."
ORIGIN
Query Match 37.5%; Score 883.6; DB 5; Length 901;
Best Local Similarity 98.9%; Pred. No. 1.3e-222;
Matches 889; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 760 GGAAGTGTCTGGCGCTGGTTTGGCTTTGCCACATGTGCCACCTCTTCTTCAATCTCC 819
Db 1 GGAAGTGTCTGGCGCTGGTTTGGCTTTGCCACATGTGCCACCTCTTCTTCAATCTCC 60
QY 820 GGAAGCAGTATCTGCAGCGCAGCAGCGCTCGCTCAAGCAGATGCAGGAGAGTTC 879
Db 61 GGAAGCAGTATCTGCAGCGCAGCAGCGCTCGCTCAAGCAGATGCAGGAGAGTTC 120
QY 880 AGGAGCATGAGGCCCGAGCTGCTGAGCGCGAGCCCAAGCCTGAGGACAGGAGTCTGAAGA 939
Db 121 AGGAGCATGAGGCCCGAGCTGCTGAGCGCGAGCCCAAGCCTGAGGACAGGAGTCTGAAGA 180
QY 940 GCGCCTGTGTAGTGTCTGAGCAGCTTCAAGTCTCGCTTCTTCTGGAGTGGGCACG 999
Db 181 GCGCCTGTGTAGTGTCTGAGCAGCTTCAAGTCTCGCTTCTTCTGGAGTGGGCACG 240
QY 1000 TTGTGTCCTGCACCGAGTGTCTACCGCGCTTGGCAGCGCCCAAGAGTGCCTATCTGCA 1059
Db 241 TTGTGTCCTGCACCGAGTGTCTACCGCGCTTGGCAGCGCCCAAGAGTGCCTATCTGCA 300
QY 1060 GACAGCGATACCCCGGATGATCCCTGTACACAGCTAATAGTTGGAGCGGCACAG 1119
```

```
Db 301 GACAGGCGATCACCCGGGTGATACCCCTGTACAAACAGCTAATAGTTTGGAAAGCGGCACAG 360
QY 1120 CTTGACCTGGAAGCACCCTGCTGCCCCCTTTTCAGGATTTTTCCTCAGAGCCCTTTGGAGG 1179
Db 361 CTTGACCTGGAAGCACCCTGCTGCCCCCTTTTCAGGATTTTTCCTCAGAGCCCTTTGGAGG 420
QY 1180 AGCAGTGGTGGGGGTAGCTGTACCTCCAGGTATGATTGAGGAGGAAATGGGTGAAGAAC 1239
Db 421 AGCAGTGGTGGGGGTAGCTGTACCTCCAGGTATGATTGAGGAGGAAATGGGTGAAGAAC 480
QY 1240 TCTCCAGACCCATGCTCCAAATGGCAGGATGTGCTTTCCCACTGAGAGGAGACCTTG 1299
Db 481 TCTCCAGACCCATGCTCCAAATGGCAGGATGTGCTTTCCCACTGAGAGGAGACCTTG 540
QY 1300 TCCATGTGAGCCTCATCAGAGCCTCACCTGGGAGGATGCGTGGCGTCTCTCCCCAGG 1359
Db 541 TCCATGTGAGCCTCATCAGAGCCTCACCTGGGAGGATGCGTGGCGTCTCTCCCCAGG 600
QY 1360 AGCCAGATCAGTGGCAGTGTGACTGAAAATGCTCATCTTAAGCACCACCAAGCCAGTGA 1419
Db 601 AGCCAGATCAGTGGCAGTGTGACTGAAAATGCTCATCTTAAGCACCACCAAGCCAGTGA 660
QY 1420 TCAGCAGCTCTTCTGTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1479
Db 661 TCAGCAGCTCTTCTGTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
QY 1480 GACTTGGTGGAGACTCAGAGGGGAGGAAAGGCTGGGCGGAGTACAGGATGAGGATGCTTG 1539
Db 721 GACTTGGTGGAGACTCAGAGGGGAGGAAAGGCTGGGCGGAGTACAGGATGAGGATGCTTG 780
QY 1540 GGTGCTGCTCCGAAAGAGACTCTGCCGAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1599
Db 781 GGTGCTGCTCCGAAAGAGACTCTGCCGAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
QY 1600 ACAGTCTTTCTTCAGAAATGTGTCAGGCTGGGCGAGTCAACTGTGTCTCTCTCTCTCTCT 1659
Db 841 ACAGTCTTTCTTCAGAAATGTGTCAGGCTGGGCGAGTCAACTGTGTCTCTCTCTCTCTCTCT 900

RESULT 13
BU528500
LOCUS
DEFINITION
AGENCOURT 6709474 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5750555
5', mRNA sequence.
ACCESSION
BU528500
VERSION
BU528500.1 GI:19370929
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1100)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM12781 row: 1 column: 12
High quality sequence stop: 681.
Location/Qualifiers
1..1100
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5750555"
/lab_host="DH10B"
FEATURES
source
```

/clone_lib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site.1: NotI; Site.2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 37.1%; Score 874.4; DB 4; Length 1100;
Best Local Similarity 94.1%; Pred. No. 3.9e-220;
Matches 964; Conservative 0; Mismatches 51; Indels 9; Gaps 5;

QY 1 ATCTTTGGCGCCACAGTCGCGCCACCGGGCTCGCCGCGTCATGAGAGCGAGGCGGC 60
DB ATCTTTGGCGCCACAGTCGCGCCACCGGGCTCGCCGCGTCATGAGAGCGAGGCGGC 134
QY 61 CCTCGTGTGCAGTTCATCTCTCTGGGACACACCTCTGTGTGTCACCGCGCCCTGTACT 120
DB CCTCGTGTGCAGTTCATCTCTCTGGGACACACCTCTGTGTGTCACCGCGCCCTGTACT 194
QY 121 CGTGTACCGGCAGAGGCGCGGCTCTCCACAGAGCTCAAGGAGCTAAAGATTCATT 180
DB CGTGTACCGGCAGAGGCGCGGCTCTCCACAGAGCTCAAGGAGCTAAAGATTCATT 254
QY 181 TGGGTGAAGATTTAAAGATTTCTTTTCAAGAGCTCCAGGAAATCGTGCCTTATGCTG 240
DB TGGGTGAAGATTTAAAGATTTCTTTTCAAGAGCTCCAGGAAATCGTGCCTTATGCTG 314
QY 241 TTATAGAAGAGCTGTGCGGTCTGTTTAAAGAAACGCTTAAAGCCAGTTTGGAAACT 300
DB TTATAGAAGAGCTGTGCGGTCTGTTTAAAGAAACGCTTAAAGCCAGTTTGGAAACT 374
QY 301 GAAGGGGTAAATTCAGCGCTGACACTTCAGGAGCACAGATGGTGTGATTCGAACCA 360
DB GAAGGGGTAAATTCAGCGCTGACACTTCAGGAGCACAGATGGTGTGATTCGAACCA 434
QY 361 CCCACCTTTGGAATGATTCCTCAAGATCATTCATCAGAGGACCAACACAGTGCCTTTG 420
DB CCCACCTTTGGAATGATTCCTCAAGATCATTCATCAGAGGACCAACACAGTGCCTTTG 494
QY 421 ACTGTGTCCCAAGAGATGGGTGGATGTGGCTGTGAGTGTGATGAGAGCCCTTGACT 480
DB ACTGTGTCCCAAGAGATGGGTGGATGTGGCTGTGAGTGTGATGAGAGCCCTTGACT 554
QY 481 CAGTGGATCTGGTCTAGAGACTGTGTATGAGAAGTTCACCCCTCAGTTCAGTCCCTCA 540
DB CAGTGGATCTGGTCTAGAGACTGTGTATGAGAAGTTCACCCCTCAGTTCAGTCCCTCA 614
QY 541 CCGATGTCTACGCGCCACTACATCAGCGGTGAGCGGCCAAAGCATCCAGAGACCGAGG 600
DB CCGATGTCTACGCGCCACTACATCAGCGGTGAGCGGCCAAAGCATCCAGAGACCGAGG 674
QY 601 AGATGCTGAAGTGGGGGACCTCTCACAGGGTGGGAACTGTGTCCTGGACAACT 660
DB AGATGCTGAAGTGGGGGACCTCTCACAGGGTGGGAACTGTGTCCTGGACAACT 734
QY 661 CTGTCCGCTGAGCGCCGCAACAGGAGTTCAGTACTATCTTAAGCAGCCAGGACTTCG 720
DB CTGTCCGCTGAGCGCCGCAACAGGAGTTCAGTACTATCTTAAGCAGCCAGGACTTCG 794
QY 721 ACAGCTGTCTGAGAGCGGAGTTCAGCGGTCTGGAAGTGTCTGGGCTGCTGTT 780
DB ACAGCTGTCTGAGAGCGGAGTTCAGCGGTCTGGAAGTGTCTGGGCTGCTGTT 854
QY 781 TTGGCTTTGACATGTGCCACCTCTTCTTCA-TTCTCCGGAAGCAGTATCTGACAGCG 839
DB TTGGCTTTGACATGTGCCACCTCTTCTTCA-TTCTCCGGAAGCAGTATCTGACAGCG 914

QY 840 CAGGAGCGCTG-CGCTCAGCAGATGCAGG-AGGATTCAGGAGCATGAGGCCAGC 897
DB CAGGAGCGCTGCGCGCTCAAGCAATGCCGAAGGAGTTTCAAGAAAGCATGAGGCCAC 974
QY 898 TGTGTAGCGGAGCAGGAGCC- TGAGGACAGGAGAGTCTGAAGAGCGCTGTGTAGTGTG 955
DB TGTGTAGCGGAGCAGGAGCCCTGAGAACAGGAGAAATTTGAAATAATGCTGTGTATGGG 1034
QY 956 TGTGAGCAGCTCAAGTCCT- ---CGTCTTTCTGAGTGTGGCAGCTTTGTTCCTGCA 1011
DB TGTGAGCAGCTTTAAGCCTGCGCGCTTTCTTGAAGTGGGGCGCTTTGTTCTCTCAC 1094
QY 1012 CCGA 1015
DB 1095 CCAA 1098

RESULT 14
BX334124
LOCUS
DEFINITION
BX334124 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD007YL23 5-PRIME, mRNA sequence.
ACCESSION
VERSION
BX334124.2 GI:46266395
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1039)
Li, W.B., Gruber, C., Jessee, J., and Pollay, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On May 1, 2003 this sequence version replaced gi:30312282.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
670.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DD007Cf12QPl&c=670.r.
FEATURES
Location/Qualifiers
1..1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD007YL23"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 36.4%; Score 858; DB 5; Length 1039;
Best Local Similarity 88.7%; Pred. No. 8.6e-216;
Matches 865; Conservative 75; Mismatches 32; Indels 3; Gaps 3;

QY 1 ATCTTTGGCGCCACAGTCGCGCCACCGGGCTCGCCGCGTCATGAGAGCGAGGCGGC 60
DB AGCCGKGGGCCACAGKSGGCSACCGGCGCKCGCCGCKCAKAGAGAGCGAGGCGGC 126
QY 61 CTCTGCTGTGCAGTTTCCTCTCTGGGACACACTCTGTGTGTGTCACCGCGCCCTGTACT 120
DB CCKCGKGGCCAGKGAACCGCGCGGGCACCCCKKGGGKCAACCGCGCCCTGTACT 186
QY 121 CGGTGTACCGGAGAGGCGCGGCTCTCCAGAGCTCAAGGAGCTAAAGATTCATT 180

```
187 CCGTGTACCGCAGAGGCCCGGTCTCCCAAGAGCKCAAGGCGCAARRAGTTCATT 246
181 TGGGTGAAGATTAAAGAGTATTTCTTTCAGAACTCCAGGAAAATCGTGCTTATGCTG 240
247 KGGGKAAGAKKAAAGAGKATCTTTTCAAGCGCCAGGAAAAGKCGKCTGAKGCGG 306
241 TTATAGAAGAGCTGTGCGTCTGTTTAAAGAAACGCTTAACAGCCAGTTTGTGGAAC 300
307 TGATAAGAGGCGGTGCGNKKGTGAAGAAACGCTTAAACGCCAGTGTGTGGAAC 366
301 GCAAGGGGTAAATTCAGCGGCTGACACTTCAGGAGCACAAAGATGTTGTTGAATCGAACCA 360
367 GCAAGGGGKAAATTCAGCGGCTGACACTTCAGGAGCACAAAGAGKGGTGTGAATCGAACCA 426
361 CCCACCTTTTGGATATGATGCTTAAAGATCATTCATCAGAGGACAAACAGTGCCTTTG 420
427 CCCACCTTTTGGATATGATGCTTAAAGATCATTCATCAGAGGACAAACAGTGCCTTTG 486
421 ACCTGTGTGCCCCAGAGGATGCGGTGATGCTGCTGTGCGAGTGTGAGCCCTCGACT 480
487 ACCTGTGTGCCCCAGAGGATGCGGTGATGCTGCTGTGCGAGTGTGAGCCCTCGACT 546
481 CAGTGTATCTGGTCTAGAGACTGTGTATGAGAAAGTTCCACCCCTCGATTTCCTTCA 540
547 CAGTGTATCTGGTCTAGAGACTGTGTATGAGAAAGTTCCACCCCTCGATTTCCTTCA 606
541 CCGATGTATCGGCCACTATCATCAGCGGTGAGCGGCCAAAGGATCCAAAGAACCGAGG 600
607 CCGATGTATCGGCCACTATCATCAGCGGTGAGC-GSCCAAAGGATCCAAAGAACCGAGG 665
601 AGATGTGAAGTGTGGGCGCCACCTCATCAGGGGTTGGCGAACTGGTCTCGACAAACAACT 660
666 AGATGTGAAGTGTGGGCGCCACCTCATCAGGGGTTGGCGAACTGGTCTCGACAAACAACT 725
661 CTGTCCGCTGTGAGCGCCGCAACAAAGGATGAGTACTATCTAAGCAGCAGGACTTCG 720
726 CTGTCCGCTGTGAGCGCCGCAACAAAGGATGAGTACTATCTAAGCAGCAGGACTTCG 785
721 ACAGCTGTGAGAGGAGGAGTGTGAGCGTCAAGGCTCTGGAAGTGTGCGCGCTGTTT 780
786 ACAGCTGTGAGAGGAGGAGTGTGAGCGTCAAGGCTCTGGAAGTGTGCGCGCTGTTT 845
781 TTGGCTTTGACATGTGCCACCTCTTCTTCACTTCCGGAAGCATATCTCAG-CGG 839
846 TTGGCTTTGACATGTGCCACCTCTTCTTCACTTCCGGAAGCATATCTCAG-CGG 905
840 CAGGAGCGCTGCGCTCAAGCAGATGAGAGGAGTTCCAGGAGCATGAGGCCAGCTG 899
906 CAGGCGSCCTGSCCTTAAGCAGATGAGAGGAGTTCCAGGAGCATGAGGCCAG-CYG 964
900 CTGAGCGGAGCGCTGAGGAGGAGTGTGAGAGCGCTGTGAGTGTGCTG 959
965 CYGAGCGSAGCCAACTGAGRSRGRGAGTTTAAAGAGCGCCCTTTTGTGTGTTT 1024
960 AGCAGCTTCAAGTCC 974
1025 AMMCCYTCAAATCC 1039
```

```
RESULT 15
AL530521
LOCUS
DEFINITION
AL530521 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD007YG15 5-PRIME, mRNA sequence.
ACCESSION
VERSION
AL530521.3
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1039)
```

AUTHORS
TITLE
JOURNAL
COMMENT

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:31069354.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
670.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?S=CS0DD007AD080P1&c=670.r.

FEATURES
source

Location/Qualifiers
1..1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD007YG15"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 36.3%; Score 856.2; DB 1; Length 1039;
Best Local Similarity 92.9%; Pred. No. 2.6e-215;
Matches 891; Conservative 23; Mismatches 42; Indels 3; Gaps 2;
QY 1 ATCTTTGGGCCACAGTCGGCCACCGGGGCTCCGCCCGCTCATGAGAGCGAGGCGGC 60
Db 41 ATCTTTGGGCCACAGTCGGCCACCGGGGCTCCGCCCGCTCATGAGAGCGAGGCGGC 100
QY 61 CCTCGTGTGCCAGTTTATCTCTCTGGGACCACTCTGTGGTTCAGCGCGCTGTACT 120
Db 101 CCTCGTGTGCCAGTTTATCTCTCTGGGACCACTCTGTGGTTCAGCGCGCTGTACT 160
QY 121 CCGTGTACCGGAGAGGCGCGGCTCTCCCAAGAGCTCAAGGAGCTTAAAAAGTTTCA 180
Db 161 CCGTGTACCGGAGAGGCGCGGCTCTCCCAAGAGCTCAAGGAGCTTAAAAAGTTTCA 220
QY 181 TGGGTGAAGATTAAAGAGTATTTCTTCAAGAGCTCCAGGAAAATGCGTGCCTTATGCTG 240
Db 221 TGGGTGAAGATTAAAGAGTATTTCTTCAAGAGCTCCAGGAAAATGCGTGCCTTATGCTG 280
QY 241 TTATAGAAGAGCTGTGCGTCTGTTTAAAGAACGCTTAACAGCCAGTTTGTGGAAC 300
Db 281 TTATAGAAGAGCTGTGCGTCTGTTTAAAGAACGCTTAAYAGCCAGTTTGTGGAAC 340
QY 301 GCAAGGGGTAAATTCAGCGGCTGACACTTCAGGAGCACAAAGATGTTGTTGAATCGAACCA 360
Db 341 GCAAGGGGTAAATTCAGCGGCTGACACTTCAGGAGCACAAAGATGTTGTTGAATCGAACCA 400
QY 361 CCCACCTTTTGGATATGATGCTTAAAGATCATTCATCAGAGGACAAACAGTGCCTTTG 420
Db 401 CCCACCTTTTGGATATGATGCTTAAAGATCATTCATCAGAGGACAAACAGTGCCTTTG 460
QY 421 ACCTGTGTGCCCCAGAGGATGCGGTGATGCTGCTGTGCGAGTGTGAGCCCTCGACT 480
Db 461 ACCTGTGTGCCCCAGAGGATGCGGTGATGCTGCTGTGCGAGTGTGAGCCCTCGACT 520
QY 481 CAGTGTATCTGGTCTAGAGACTGTGTATGAGAAAGTTCCACCCCTCGATTTCCTTCA 540
Db 521 CAGTGTATCTGGTCTAGAGACTGTGTATGAGAAAGTTCCACCCCTCGATTTCCTTCA 580
QY 541 CCGATGTATCGGCCACTATCATCAGCGGTGAGCGGCCAAAGGATCCAAAGAACCGAGG 600
Db 581 CCGATGTATCGGCCACTATCATCAGCGGTGAGCGGCCAAAGGATCCAAAGAACCGAGG 640

```
QY 601 AGATGCTGAAGCTGGGGCCACCCTCACAGGGTTGGCGAACTGCTCTGGACAACAAC 660
Db |||||
QY 641 AGATGCTGAAGCTGGGGCCACCCTCACAGGGTTGGCGAACTGCTCTGGACAACAAC 700
Db |||||
QY 661 CTGTCCGCTGTCAGCGCGCCCAAAACAAGGCATGCAGTACTATCTAAGCAGCCAGGACTTCG 720
Db |||||
QY 701 CTGTTTCGCTGTCAGCGCGCCCAAAACAAGGCATGCAGTACTATCTAAGCAGCCAGGACTTCG 760
Db |||||
QY 721 ACAGCCTGCTGTCAGAGGCGAGGAGTCGAGCGTCAGGCTCTGGAAGGTGCTGGCGCTGGTTT 780
Db |||||
QY 761 ACAGCCTGCTGTCAGAGGCGAGGAGTCGAGCGTCAGGCTCTGGAAGGTGCTGGCGCTGGTTT 820
Db |||||
QY 781 TTGGCTTTGCCACATGTGCCACCTCTTCTTCACTTCCGGAAGCAGTATCTGAGCGGC 840
Db |||||
QY 821 TTGGCTTTGCCACATGTGCCACCTCTTCTTCACTTCCGGAAGCAGTATCTGAGCGGC 880
Db |||||
QY 841 AGGAGCGCTGCGCCTCAAGCAGATGCAGGAGGAGTTCAGGAGCATGAGGCCAGCTGC 900
Db |||||
QY 881 AGGAGCSCCTTCGCTTAAAGCAGATGCAGAGGAG-TCCAVGAGCMTKAHGGCCACCCGY 939
Db |||||
QY 901 TGAGCCGAGCCAGCCTGAGGACAGGAGAGTCTGAAGAGCGCCTGTGTAGTGTCTG 959
Db |||||
QY 940 --GACCCGACCAAGCYTGAGGMAAGAAATTTAARACGCTTTTGTGTTTTTRACCG 996
Db |||||
```

Search completed: November 8, 2004, 02:50:31
Job time : 6593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 2004, 17:38:28 ; Search time 142 Seconds
(without alignments)
875.602 Million cell updates/sec

Title: US-09-978-360A-437

Perfect score: 1831

Sequence: 1 MESGGPRLCQFILLGTTSV.....PKKPCICQAIRTVPLYS 352

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1831	100.0	352	11	US-09-978-360A-437
2	1831	100.0	352	14	US-10-024-298A-75
3	1831	100.0	352	14	US-10-042-211A-75
4	1831	100.0	352	14	US-10-315-664-105
5	1831	100.0	352	15	US-10-221-625-25
6	1831	100.0	352	15	US-10-169-395-92
7	1831	100.0	352	15	US-10-617-217A-75
8	1831	100.0	352	17	US-10-024-298A-75
9	1826	99.7	392	9	US-09-764-864-801
10	1824	99.6	352	14	US-10-024-298A-73
11	1824	99.6	352	14	US-10-042-211A-73
12	1824	99.6	352	15	US-10-617-217A-73
13	1824	99.6	352	17	US-10-024-298A-73

14	901	49.2	174	9	US-09-764-864-1262	Sequence 1262, Ap
15	854	46.6	165	15	US-10-264-237-2680	Sequence 2680, Ap
16	295	16.1	339	15	US-10-424-599-158556	Sequence 158556, Ap
17	288.5	15.8	343	17	US-10-425-115-364826	Sequence 364826, Ap
18	267.5	14.6	338	17	US-10-739-930-6185	Sequence 6185, Ap
19	262	14.3	350	17	US-10-425-115-364805	Sequence 364805, Ap
20	252.5	13.8	375	17	US-10-739-930-10614	Sequence 10614, A
21	231	12.6	332	15	US-10-424-599-224671	Sequence 224671, A
22	214	11.7	358	15	US-10-425-114-47706	Sequence 47706, A
23	213	11.6	310	16	US-10-437-963-111394	Sequence 111394, A
24	195.5	10.7	378	16	US-10-437-963-1138606	Sequence 138606, Ap
25	188.5	10.3	332	14	US-10-195-144-83	Sequence 83, Appl
26	188.5	10.3	332	14	US-10-345-072-83	Sequence 83, Appl
27	184.5	10.1	163	15	US-10-425-114-65305	Sequence 65305, A
28	183	10.0	236	15	US-10-425-114-38446	Sequence 38446, A
29	183	10.0	236	17	US-10-425-115-253264	Sequence 253264, A
30	172	9.4	366	17	US-10-739-930-7102	Sequence 7102, Ap
31	163	8.9	257	9	US-09-949-842-14	Sequence 14, Appl
32	163	8.9	403	13	US-10-087-192-114	Sequence 114, Appl
33	155.5	8.5	303	14	US-10-094-749-2475	Sequence 2475, Ap
34	155.5	8.5	696	14	US-10-094-749-2425	Sequence 2425, Ap
35	155.5	8.5	723	14	US-10-104-047-2572	Sequence 2572, Ap
36	148.5	8.1	272	15	US-10-108-260A-4608	Sequence 4608, Ap
37	145.5	7.9	336	15	US-10-451-168-89	Sequence 89, Appl
38	144.5	7.9	169	17	US-10-425-115-364837	Sequence 364837, Ap
39	144.5	7.9	438	8	US-08-464-588-2	Sequence 2, Appl
40	144.5	7.9	438	14	US-10-323-643-2	Sequence 323, Appl
41	144.5	7.9	618	14	US-10-153-668-338	Sequence 338, Appl
42	144.5	7.9	618	14	US-10-207-655-200	Sequence 200, Appl
43	144.5	7.9	618	14	US-10-232-286-2	Sequence 2, Appl
44	144.5	7.9	618	14	US-10-366-307-4	Sequence 4, Appl
45	144.5	7.9	618	15	US-10-361-270-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-978-360A-437
; Sequence 437, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.USA.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810

```
; SOFTWARE: Patent.pm
; SEQ ID NO 437
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -23..-1
US-09-978-360A-437

Query Match      100.0%; Score 1831; DB 11; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-162;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESGGPRLCQFILLGTTTSVTTAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
Db 1 MESGGPRLCQFILLGTTTSVTTAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
QY 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWNCDSKIIHOR 120
Db 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWNCDSKIIHOR 120
QY 121 TMTVPDLVPHEDGVAVRVLKPLDSVDLGLTETVYKEKHPHSIQSFTDVIHGYISGERPK 180
Db 121 TMTVPDLVPHEDGVAVRVLKPLDSVDLGLTETVYKEKHPHSIQSFTDVIHGYISGERPK 180
QY 181 GQOETEMLKVGATLTGVLGELVDNNSVRLQPPKQMGYYLSSQDFDLSLLQROESSVRLW 240
Db 181 GQOETEMLKVGATLTGVLGELVDNNSVRLQPPKQMGYYLSSQDFDLSLLQROESSVRLW 240
QY 241 KVALVFGATCATLFFILRKQYLOERLRKQMOEFOEHEAQLLSRAKPEDRESLKS 300
Db 241 KVALVFGATCATLFFILRKQYLOERLRKQMOEFOEHEAQLLSRAKPEDRESLKS 300
QY 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLYS 352
Db 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLYS 352

RESULT 2
US-10-024-298A-75
; Sequence 75, Application US/10024298A
; Publication No. US20030143540A1
; GENERAL INFORMATION:
; APPLICANT: ASAH KASEI KABUSHIKI KAISHA
; APPLICANT: AKIO MATSUDA
; APPLICANT: GOICHI HONDA
; APPLICANT: SHUJI MURAMATSU
; APPLICANT: YUKIO NAGANO
; TITLE OF INVENTION: NF-K B Activating Gene
; FILE REFERENCE: 1254-0191P
; CURRENT APPLICATION NUMBER: US/10/024,298A
; PRIOR FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: 60/314,385
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/278,641
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/258,315
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: JP254018/2001
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: JP0088912/2001
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: JP402288/2000
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-298A-75

Query Match      100.0%; Score 1831; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-162;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESGGPRLCQFILLGTTTSVTTAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
Db 1 MESGGPRLCQFILLGTTTSVTTAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
QY 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWNCDSKIIHOR 120
Db 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWNCDSKIIHOR 120
QY 121 TMTVPDLVPHEDGVAVRVLKPLDSVDLGLTETVYKEKHPHSIQSFTDVIHGYISGERPK 180
Db 121 TMTVPDLVPHEDGVAVRVLKPLDSVDLGLTETVYKEKHPHSIQSFTDVIHGYISGERPK 180
QY 181 GQOETEMLKVGATLTGVLGELVDNNSVRLQPPKQMGYYLSSQDFDLSLLQROESSVRLW 240
Db 181 GQOETEMLKVGATLTGVLGELVDNNSVRLQPPKQMGYYLSSQDFDLSLLQROESSVRLW 240
QY 241 KVALVFGATCATLFFILRKQYLOERLRKQMOEFOEHEAQLLSRAKPEDRESLKS 300
Db 241 KVALVFGATCATLFFILRKQYLOERLRKQMOEFOEHEAQLLSRAKPEDRESLKS 300
QY 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLYS 352
Db 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLYS 352

RESULT 3
US-10-042-211A-75
; Sequence 75, Application US/10042211A
; Publication No. US20030170719A1
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Akio et al.
; TITLE OF INVENTION: NFkB Activating Gene
; FILE REFERENCE: 1254-0192P
; CURRENT APPLICATION NUMBER: US/10/042,211A
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: JP 2000-402288
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: JP 2001-088912
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: JP 2001-254018
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/258,315
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/278,640
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/314,385
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-211A-75

Query Match      100.0%; Score 1831; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-162;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESGGPRLCQFILLGTTTSVTTAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
Db 1 MESGGPRLCQFILLGTTTSVTTAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
QY 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWNCDSKIIHOR 120
Db 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWNCDSKIIHOR 120
QY 121 TMTVPDLVPHEDGVAVRVLKPLDSVDLGLTETVYKEKHPHSIQSFTDVIHGYISGERPK 180
Db 121 TMTVPDLVPHEDGVAVRVLKPLDSVDLGLTETVYKEKHPHSIQSFTDVIHGYISGERPK 180
```

Db 121 TMTVPDLVPHEDGVDVAVRVLPKLDSDVLGLTVEYKFPSPISQSFDTDVIGHYISGERPK 180
QY 181 GIQTEMLKVGATLTGVELVLNNNSVRLQPPKQGMYYLSSQDFDLSLQORQESSVRLW 240
Db 181 GIQTEMLKVGATLTGVELVLNNNSVRLQPPKQGMYYLSSQDFDLSLQORQESSVRLW 240
QY 241 KVALVFGFATCATLFFTLRKQYLRQERLRLKQMEEFQHEAQLLSRAKPEDRESLKS 300
Db 241 KVALVFGFATCATLFFTLRKQYLRQERLRLKQMEEFQHEAQLLSRAKPEDRESLKS 300
QY 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352
Db 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352

RESULT 4

US-10-315-664-105
; Sequence 105, Application US/10315664
; Publication No. US20030203377A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal
; FILE REFERENCE: Peptides
; CURRENT APPLICATION NUMBER: US/10/315.664
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/599,360
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 105
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -23...-1
US-10-315-664-105

Query Match 100.0%; Score 1831; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-162;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESGGRPSLCQFILLGTTSVVTAALYSYVRQKARVSQELKAKVHLGEDLKSILSEAPG 60
Db 1 MESGGRPSLCQFILLGTTSVVTAALYSYVRQKARVSQELKAKVHLGEDLKSILSEAPG 60
QY 61 KCVPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLOEHKVMWNRTHLWNCDSKIIHQ 120
Db 61 KCVPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLOEHKVMWNRTHLWNCDSKIIHQ 120
QY 121 TMTVPDLVPHEDGVDVAVRVLPKLDSDVLGLTVEYKFPSPISQSFDTDVIGHYISGERPK 180
Db 121 TMTVPDLVPHEDGVDVAVRVLPKLDSDVLGLTVEYKFPSPISQSFDTDVIGHYISGERPK 180
QY 181 GIQTEMLKVGATLTGVELVLNNNSVRLQPPKQGMYYLSSQDFDLSLQORQESSVRLW 240
Db 181 GIQTEMLKVGATLTGVELVLNNNSVRLQPPKQGMYYLSSQDFDLSLQORQESSVRLW 240
QY 241 KVALVFGFATCATLFFTLRKQYLRQERLRLKQMEEFQHEAQLLSRAKPEDRESLKS 300
Db 241 KVALVFGFATCATLFFTLRKQYLRQERLRLKQMEEFQHEAQLLSRAKPEDRESLKS 300
QY 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352
Db 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352

Db 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352

RESULT 5

US-10-221-625-25
; Sequence 25, Application US/10221625
; Publication No. US20040033942A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: LU, Dyoung Aina M.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: MATHUR, Preete
; APPLICANT: SHAH, Purvi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; TITLE OF INVENTION: TRANSCRIPTION FACTORS
; FILE REFERENCE: PR-0761 PCT
; CURRENT APPLICATION NUMBER: US/10/221.625
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040033942A1 1383473CD1
US-10-221-625-25

Query Match 100.0%; Score 1831; DB 15; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-162;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESGGRPSLCQFILLGTTSVVTAALYSYVRQKARVSQELKAKVHLGEDLKSILSEAPG 60
Db 1 MESGGRPSLCQFILLGTTSVVTAALYSYVRQKARVSQELKAKVHLGEDLKSILSEAPG 60
QY 61 KCVPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLOEHKVMWNRTHLWNCDSKIIHQ 120
Db 61 KCVPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLOEHKVMWNRTHLWNCDSKIIHQ 120
QY 121 TMTVPDLVPHEDGVDVAVRVLPKLDSDVLGLTVEYKFPSPISQSFDTDVIGHYISGERPK 180
Db 121 TMTVPDLVPHEDGVDVAVRVLPKLDSDVLGLTVEYKFPSPISQSFDTDVIGHYISGERPK 180
QY 181 GIQTEMLKVGATLTGVELVLNNNSVRLQPPKQGMYYLSSQDFDLSLQORQESSVRLW 240
Db 181 GIQTEMLKVGATLTGVELVLNNNSVRLQPPKQGMYYLSSQDFDLSLQORQESSVRLW 240
QY 241 KVALVFGFATCATLFFTLRKQYLRQERLRLKQMEEFQHEAQLLSRAKPEDRESLKS 300
Db 241 KVALVFGFATCATLFFTLRKQYLRQERLRLKQMEEFQHEAQLLSRAKPEDRESLKS 300
QY 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352
Db 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352

RESULT 6

US-10-169-395-92
; Sequence 92, Application US/10169395
; Publication No. US20040034192A1
; GENERAL INFORMATION:
; APPLICANT: KATO, Seishi
; APPLICANT: KIMURA, Tomoko

; TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAs ENCODING
; TITLE OF INVENTION: THESE PROTEINS
; FILE REFERENCE: 01997.015100.US
; CURRENT APPLICATION NUMBER: US/10/169,395
; CURRENT FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: JP 2000-585
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-588
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-2299
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-26862
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: JP 2000-58367
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/JP00/09359
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 150
; SEQ ID NO 92
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-169-395-92

Query Match 100.0%; Score 1831; DB 15; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-162;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESGGRPSLCQFILLGTTSVVTAALYSVYRQKARVSOELKGAKKVVHLGDLKSIILSEAPG 60
DB 1 MESGGRPSLCQFILLGTTSVVTAALYSVYRQKARVSOELKGAKKVVHLGDLKSIILSEAPG 60
QY 61 KCVPYAVIEGAVRSVKETLNSQFVENCCKGVIOQLTLQEHKVMVNRTHLWDCSKIHHOR 120
DB 61 KCVPYAVIEGAVRSVKETLNSQFVENCCKGVIOQLTLQEHKVMVNRTHLWDCSKIHHOR 120
QY 121 TMTVPFDLVPHEDGVDVAVRVLPKLDSDVLDGLETVYVEKPHPSIQSFTDVIHGIISGERPK 180
DB 121 TMTVPFDLVPHEDGVDVAVRVLPKLDSDVLDGLETVYVEKPHPSIQSFTDVIHGIISGERPK 180
QY 181 GQOETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQSDPDSLLQEQESSVRLW 240
DB 181 GQOETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQSDPDSLLQEQESSVRLW 240
QY 241 KVALVFGFATCATLFFILRKQYLQORERLRKQOEFEQHEAQLLSRAKPEDRESLKS 300
DB 241 KVALVFGFATCATLFFILRKQYLQORERLRKQOEFEQHEAQLLSRAKPEDRESLKS 300
QY 301 ACVCLSPKSCVFLECGHVCSTCYRALPEPKKPCICRQAITRVIPLYS 352
DB 301 ACVCLSPKSCVFLECGHVCSTCYRALPEPKKPCICRQAITRVIPLYS 352

RESULT 7

US-10-617-217A-75
; Sequence 75, Application US/10617217A
; Publication No. US2004008196A1
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Akio et al.
; TITLE OF INVENTION: NF-KB ACTIVATING GENE
; FILE REFERENCE: 1254-0229P
; CURRENT APPLICATION NUMBER: US/10/617,217A
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: JP 2000-402288
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: JP 2001-089912
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: JP 2001-254018
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/258,315
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/278,640
; PRIOR FILING DATE: 2001-03-26

; PRIOR APPLICATION NUMBER: US 60/314,385
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 75
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-617-217A-75

Query Match 100.0%; Score 1831; DB 15; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-162;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESGGRPSLCQFILLGTTSVVTAALYSVYRQKARVSOELKGAKKVVHLGDLKSIILSEAPG 60
DB 1 MESGGRPSLCQFILLGTTSVVTAALYSVYRQKARVSOELKGAKKVVHLGDLKSIILSEAPG 60
QY 61 KCVPYAVIEGAVRSVKETLNSQFVENCCKGVIOQLTLQEHKVMVNRTHLWDCSKIHHOR 120
DB 61 KCVPYAVIEGAVRSVKETLNSQFVENCCKGVIOQLTLQEHKVMVNRTHLWDCSKIHHOR 120
QY 121 TMTVPFDLVPHEDGVDVAVRVLPKLDSDVLDGLETVYVEKPHPSIQSFTDVIHGIISGERPK 180
DB 121 TMTVPFDLVPHEDGVDVAVRVLPKLDSDVLDGLETVYVEKPHPSIQSFTDVIHGIISGERPK 180
QY 181 GQOETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQSDPDSLLQEQESSVRLW 240
DB 181 GQOETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQSDPDSLLQEQESSVRLW 240
QY 241 KVALVFGFATCATLFFILRKQYLQORERLRKQOEFEQHEAQLLSRAKPEDRESLKS 300
DB 241 KVALVFGFATCATLFFILRKQYLQORERLRKQOEFEQHEAQLLSRAKPEDRESLKS 300
QY 301 ACVCLSPKSCVFLECGHVCSTCYRALPEPKKPCICRQAITRVIPLYS 352
DB 301 ACVCLSPKSCVFLECGHVCSTCYRALPEPKKPCICRQAITRVIPLYS 352

RESULT 8

US-10-024-298A-75
; Sequence 75, Application US/10024298A
; Publication No. US20040214167A9
; GENERAL INFORMATION:
; APPLICANT: ASAH KASEI KABUSHIKI KAISHA
; APPLICANT: AKIO MATSUDA
; APPLICANT: GOICHI HONDA
; APPLICANT: SHUJI MURAMATSU
; APPLICANT: YUKIKO NAGANO
; TITLE OF INVENTION: NF-KB Activating Gene
; FILE REFERENCE: 1254-0191P
; CURRENT APPLICATION NUMBER: US/10/024,298A
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: 60/314,385
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/278,641
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/258,315
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: JP254018/2001
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: JP0088912/2001
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: JP402288/2000
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 75
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-298A-75

Query Match 100.0%; Score 1831; DB 17; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-162;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESGGRPSLCQFILLGTTTAAALYSVYRKARVSQELKGAKVHLGEDLKSILSEAPG 60
DB 1 MESGGRPSLCQFILLGTTTAAALYSVYRKARVSQELKGAKVHLGEDLKSILSEAPG 60

QY 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKI IHR 120
DB 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKI IHR 120

QY 121 TNPVPDLVPHEDGVDVAVRVLKPLDSVDLGLTVEYKHPHSIQSFTDVI GHYISGERPK 180
DB 121 TNPVPDLVPHEDGVDVAVRVLKPLDSVDLGLTVEYKHPHSIQSFTDVI GHYISGERPK 180

QY 181 GIQTEEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQROESSVRLW 240
DB 181 GIQTEEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQROESSVRLW 240

QY 241 KVLALVFGFATCATLFFILRKQYLRQERLRLKQMOEFQEHQAQLLSRAKPEDRESLSKS 300
DB 241 KVLALVFGFATCATLFFILRKQYLRQERLRLKQMOEFQEHQAQLLSRAKPEDRESLSKS 300

QY 301 ACVCLSSFKSCVFLGCHVCSCTECYRALPEPKKPCICRQAITRVIPLYS 352
DB 301 ACVCLSSFKSCVFLGCHVCSCTECYRALPEPKKPCICRQAITRVIPLYS 352

RESULT 9

US-09-764-864-801
; Sequence 801, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 801
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (238)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-801

Query Match 99.7%; Score 1826; DB 9; Length 392;
Best Local Similarity 99.7%; Pred. No. 5.5e-162;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESGGRPSLCQFILLGTTTAAALYSVYRKARVSQELKGAKVHLGEDLKSILSEAPG 60
DB 41 MESGGRPSLCQFILLGTTTAAALYSVYRKARVSQELKGAKVHLGEDLKSILSEAPG 100

QY 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKI IHR 120
DB 101 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKI IHR 160

QY 121 TNPVPDLVPHEDGVDVAVRVLKPLDSVDLGLTVEYKHPHSIQSFTDVI GHYISGERPK 180
DB 161 TNPVPDLVPHEDGVDVAVRVLKPLDSVDLGLTVEYKHPHSIQSFTDVI GHYISGERPK 220

QY 181 GIQTEEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQROESSVRLW 240
DB 221 GIQTEEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQROESSVRLW 280

QY 241 KVLALVFGFATCATLFFILRKQYLRQERLRLKQMOEFQEHQAQLLSRAKPEDRESLSKS 300

Db 281 KVLALVFGFATCATLFFILRKQYLRQERLRLKQMOEFQEHQAQLLSRAKPEDRESLSKS 340
QY 301 ACVCLSSFKSCVFLGCHVCSCTECYRALPEPKKPCICRQAITRVIPLYS 352
Db 341 ACVCLSSFKSCVFLGCHVCSCTECYRALPEPKKPCICRQAITRVIPLYS 392

RESULT 10

US-10-024-298A-73
; Sequence 73, Application US/10024298A
; Publication No. US20030143540A1
; GENERAL INFORMATION:
; APPLICANT: ASAH KASEI KABUSHIKI KAISHA
; APPLICANT: Goichi HONDA
; APPLICANT: Shuji MURAMATSU
; APPLICANT: Yukiko NAGANO
; TITLE OF INVENTION: NF-K B Activating Gene
; FILE REFERENCE: 1254-0191P
; CURRENT APPLICATION NUMBER: US/10/024,298A
; PRIOR FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: 60/314,385
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/278,641
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/258,315
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: JP254018/2001
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: JP008912/2001
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: JP402288/2000
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-298A-73

Query Match 99.6%; Score 1824; DB 14; Length 352;
Best Local Similarity 99.7%; Pred. No. 7.3e-162;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESGGRPSLCQFILLGTTTAAALYSVYRKARVSQELKGAKVHLGEDLKSILSEAPG 60
DB 1 MESGGRPSLCQFILLGTTTAAALYSVYRKARVSQELKGAKVHLGEDLKSILSEAPG 60

QY 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKI IHR 120
DB 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKI IHR 120

QY 121 TNPVPDLVPHEDGVDVAVRVLKPLDSVDLGLTVEYKHPHSIQSFTDVI GHYISGERPK 180
DB 121 TNPVPDLVPHEDGVDVAVRVLKPLDSVDLGLTVEYKHPHSIQSFTDVI GHYISGERPK 180

QY 181 GIQTEEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQROESSVRLW 240
DB 181 GIQTEEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQROESSVRLW 240

QY 241 KVLALVFGFATCATLFFILRKQYLRQERLRLKQMOEFQEHQAQLLSRAKPEDRESLSKS 300
DB 241 KVLALVFGFATCATLFFILRKQYLRQERLRLKQMOEFQEHQAQLLSRAKPEDRESLSKS 300

QY 301 ACVCLSSFKSCVFLGCHVCSCTECYRALPEPKKPCICRQAITRVIPLYS 352
DB 301 ACVCLSSFKSCVFLGCHVCSCTECYRALPEPKKPCICRQAITRVIPLYS 352

RESULT 11

US-10-042-211A-73

; Sequence 73, Application US/10042211A	
; Publication No. US2003010719A1	
; GENERAL INFORMATION:	
; APPLICANT: MATSUDA, Akio et al.	
; TITLE OF INVENTION: NF-KB Activating Gene	
; FILE REFERENCE: 1254-0192P	
; CURRENT APPLICATION NUMBER: US/10/042,211A	
; CURRENT FILING DATE: 2002-01-11	
; PRIOR APPLICATION NUMBER: JP 2000-402288	
; PRIOR FILING DATE: 2000-12-28	
; PRIOR APPLICATION NUMBER: JP 2001-088912	
; PRIOR FILING DATE: 2001-03-26	
; PRIOR APPLICATION NUMBER: JP 2001-254018	
; PRIOR FILING DATE: 2001-08-24	
; PRIOR APPLICATION NUMBER: US 60/258,315	
; PRIOR FILING DATE: 2000-12-28	
; PRIOR APPLICATION NUMBER: US 60/278,640	
; PRIOR FILING DATE: 2001-03-26	
; PRIOR APPLICATION NUMBER: US 60/314,385	
; PRIOR FILING DATE: 2001-08-24	
; NUMBER OF SEQ ID NOS: 182	
; SOFTWARE: PatentIn Ver. 2.0	
; SEQ ID NO 73	
; LENGTH: 352	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-10-042-211A-73	
Query Match 99.6%; Score 1824; DB 14; Length 352;	
Best Local Similarity 99.7%; Pred. No. 7.3e-162;	
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 MESGGPRLCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
Db	1 MESGGPRLCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
QY	61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWDCSKIIHOR 120
Db	61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWDCSKIIHOR 120
QY	121 TMTVPDLVPHEGVDVAVRVLKPLDSVDLGLTETVYKEKHPSPQSFTDVIHGYISGERPK 180
Db	121 TMTVPDLVPHEGVDVAVRVLKPLDSVDLGLTETVYKEKHPSPQSFTDVIHGYISGERPK 180
QY	181 GIQETEMLKVGATLTGVLGELVDLNNNSVRLQPPKQGMYYLSSQDFDLSLLQROESSVRLW 240
Db	181 GIQETEMLKVGATLTGVLGELVDLNNNSVRLQPPKQGMYYLSSQDFDLSLLQROESSVRLW 240
QY	241 KVLALVFGFATCATLFFILRKQYLQQRERLRLKQMOEFOEHEAQLLSRAKPEDRESLKS 300
Db	241 KVLALVFGFATCATLFFILRKQYLQQRERLRLKQMOEFOEHEAQLLSRAKPEDRESLKS 300
QY	301 ACVVCLSSPKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLVNS 352
Db	301 ACVVCLSSPKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLVNS 352
RESULT 12	
US-10-617-217A-73	
; Sequence 73, Application US/10617217A	
; Publication No. US20040081986A1	
; GENERAL INFORMATION:	
; APPLICANT: MATSUDA, Akio et al.	
; TITLE OF INVENTION: NF-KB ACTIVATING GENE	
; FILE REFERENCE: 1254-0229P	
; CURRENT APPLICATION NUMBER: US/10/617,217A	
; CURRENT FILING DATE: 2003-07-11	
; PRIOR APPLICATION NUMBER: JP 2000-402288	
; PRIOR FILING DATE: 2000-12-28	
; PRIOR APPLICATION NUMBER: JP 2001-088912	
; PRIOR FILING DATE: 2001-03-26	
; PRIOR APPLICATION NUMBER: JP 2001-254018	
; PRIOR FILING DATE: 2001-08-24	
; SEQ ID NO 73	
; LENGTH: 352	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-10-042-211A-73	
Query Match 99.6%; Score 1824; DB 14; Length 352;	
Best Local Similarity 99.7%; Pred. No. 7.3e-162;	
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 MESGGPRLCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
Db	1 MESGGPRLCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
QY	61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWDCSKIIHOR 120
Db	61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWDCSKIIHOR 120
QY	121 TMTVPDLVPHEGVDVAVRVLKPLDSVDLGLTETVYKEKHPSPQSFTDVIHGYISGERPK 180
Db	121 TMTVPDLVPHEGVDVAVRVLKPLDSVDLGLTETVYKEKHPSPQSFTDVIHGYISGERPK 180
QY	181 GIQETEMLKVGATLTGVLGELVDLNNNSVRLQPPKQGMYYLSSQDFDLSLLQROESSVRLW 240
Db	181 GIQETEMLKVGATLTGVLGELVDLNNNSVRLQPPKQGMYYLSSQDFDLSLLQROESSVRLW 240
QY	241 KVLALVFGFATCATLFFILRKQYLQQRERLRLKQMOEFOEHEAQLLSRAKPEDRESLKS 300
Db	241 KVLALVFGFATCATLFFILRKQYLQQRERLRLKQMOEFOEHEAQLLSRAKPEDRESLKS 300
QY	301 ACVVCLSSPKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLVNS 352
Db	301 ACVVCLSSPKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLVNS 352
RESULT 13	
US-10-024-298A-73	
; Sequence 73, Application US/10024298A	
; Publication No. US20040214167A9	
; GENERAL INFORMATION:	
; APPLICANT: ASAHU KASEI KABUSHIKI KAISHA	
; APPLICANT: AKIO MATSUDA	
; APPLICANT: GOICHI HONDA	
; APPLICANT: SHUJI MURAMATSU	
; APPLICANT: YUKIKO NAGANO	
; TITLE OF INVENTION: NF-K B Activating Gene	
; FILE REFERENCE: 1254-0191P	
; CURRENT APPLICATION NUMBER: US/10/024,298A	
; CURRENT FILING DATE: 2003-04-08	
; PRIOR APPLICATION NUMBER: 60/314,385	
; PRIOR FILING DATE: 2001-08-24	
; PRIOR APPLICATION NUMBER: 60/278,641	
; PRIOR FILING DATE: 2001-03-26	
; PRIOR APPLICATION NUMBER: 60/258,315	
; PRIOR FILING DATE: 2000-12-28	
; PRIOR APPLICATION NUMBER: JP254018/2001	
; PRIOR FILING DATE: 2001-08-24	
; PRIOR APPLICATION NUMBER: JP0088912/2001	
; PRIOR FILING DATE: 2001-03-26	
; PRIOR APPLICATION NUMBER: JP402288/2000	
; PRIOR FILING DATE: 2000-12-28	
; NUMBER OF SEQ ID NOS: 182	
; SOFTWARE: PatentIn Ver. 2.0	
; SEQ ID NO 73	
; LENGTH: 352	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-10-617-217A-73	
Query Match 99.6%; Score 1824; DB 15; Length 352;	
Best Local Similarity 99.7%; Pred. No. 7.3e-162;	
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 MESGGPRLCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
Db	1 MESGGPRLCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKKVKHLGDLKSLSEAPG 60
QY	61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWDCSKIIHOR 120
Db	61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNRTHLWDCSKIIHOR 120
QY	121 TMTVPDLVPHEGVDVAVRVLKPLDSVDLGLTETVYKEKHPSPQSFTDVIHGYISGERPK 180
Db	121 TMTVPDLVPHEGVDVAVRVLKPLDSVDLGLTETVYKEKHPSPQSFTDVIHGYISGERPK 180
QY	181 GIQETEMLKVGATLTGVLGELVDLNNNSVRLQPPKQGMYYLSSQDFDLSLLQROESSVRLW 240
Db	181 GIQETEMLKVGATLTGVLGELVDLNNNSVRLQPPKQGMYYLSSQDFDLSLLQROESSVRLW 240
QY	241 KVLALVFGFATCATLFFILRKQYLQQRERLRLKQMOEFOEHEAQLLSRAKPEDRESLKS 300
Db	241 KVLALVFGFATCATLFFILRKQYLQQRERLRLKQMOEFOEHEAQLLSRAKPEDRESLKS 300
QY	301 ACVVCLSSPKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLVNS 352
Db	301 ACVVCLSSPKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVIPLVNS 352


```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-298A-73

Query Match          99.6%; Score 1824; DB 17; Length 352;
Best Local Similarity 99.7%; Pred. No. 7.3e-162;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESGGRPSCQFILLGTTSTVTAALSVYRKARVSOELKGAKKVLHGEDLKSLSEAPG 60
Db 1 MESGGRPSCQFILLGTTSTVTAALSVYRKARVSOELKGAKKVLHGEDLKSLSEAPG 60
QY 61 KCPYAVIEGAVRSVKETLNSQFVENCKVGIQRLTLQEHKVMWNRTHLWDCSKLIHOR 120
Db 61 KCPYAVIEGAVRSVKETLNSQFVENCKVGIQRLTLQEHKVMWNRTHLWDCSKLIHOR 120
QY 121 TMTVPDLVPHEDGVAVRVLKPLSDVLGLTVTEKHPSTQSTDTDVIGHYISGERPK 180
Db 121 TMTVPDLVPHEDGVAVRVLKPLSDVLGLTVTEKHPSTQSTDTDVIGHYISGERPK 180
QY 181 GIQTEEMLKVGATLTGVELVDNNSVRLQPPKQGMQYLLSSQDFDSLQEQESSVRLW 240
Db 181 GIQTEEMLKVGATLTGVELVDNNSVRLQPPKQGMQYLLSSQDFDSLQEQESSVRLW 240
QY 241 KVLALVFGFATCATLFFILRKQYLRQERLRKQMOEEFQEHQAQLLSRAKPEDRESLKS 300
Db 241 KVLALVFGFATCATLFFILRKQYLRQERLRKQMOEEFQEHQAQLLSRAKPEDRESLKS 300
QY 301 ACWCLSSFKSCVFLGCHGVCSTECYRALPPEPKKPCICRQAITRVIPLYS 352
Db 301 ACWCLSSFKSCVFLGCHGVCSTECYRALPPEPKKPCICRQAITRVIPLYS 352

RESULT 14
US-09-764-864-1262
; Sequence 1262, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1262
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (117)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1262

Query Match          49.2%; Score 901; DB 9; Length 174;
Best Local Similarity 99.4%; Pred. No. 6.7e-76;
Matches 173; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 179 PKGIQTEEMLKVGATLTGVELVDNNSVRLQPPKQGMQYLLSSQDFDSLQEQESSVR 238
Db 1 PKGIQTEEMLKVGATLTGVELVDNNSVRLQPPKQGMQYLLSSQDFDSLQEQESSVR 60
QY 239 LKVLALVFGFATCATLFFILRKQYLRQERLRKQMOEEFQEHQAQLLSRAKPEDRESL 298
Db 61 LKVLALVFGFATCATLFFILRKQYLRQERLRKQMOEEFQEHQAQLLSRAKPEDRESL 120
QY 299 KSACVCLSSFKSCVFLGCHGVCSTECYRALPPEPKKPCICRQAITRVIPLYS 352
Db 121 KSACVCLSSFKSCVFLGCHGVCSTECYRALPPEPKKPCICRQAITRVIPLYS 174
```

```
RESULT 15
US-10-264-237-2680
; Sequence 2680, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2680
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (108)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-2680

Query Match          46.6%; Score 854; DB 15; Length 165;
Best Local Similarity 99.4%; Pred. No. 1.6e-71;
Matches 164; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 MLKVGATLTGVELVDNNSVRLQPPKQGMQYLLSSQDFDSLQEQESSVRLWKVLALVF 247
Db 1 MLKVGATLTGVELVDNNSVRLQPPKQGMQYLLSSQDFDSLQEQESSVRLWKVLALVF 60
QY 248 GPATCATLFFILRKQYLRQERLRKQMOEEFQEHQAQLLSRAKPEDRESLKSACVCL 307
Db 61 GPATCATLFFILRKQYLRQERLRKQMOEEFQEHQAQLLSRAKPEDRESLKSACVCL 120
QY 308 SPKSCVFLGCHGVCSTECYRALPPEPKKPCICRQAITRVIPLYS 352
Db 121 SPKSCVFLGCHGVCSTECYRALPPEPKKPCICRQAITRVIPLYS 165
```

Search completed: November 6, 2004, 17:53:27
Job time : 144 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 2004, 17:24:18 ; Search time 155 Seconds

(without alignments)
814.663 Million cell updates/sec

Title: US-09-978-360A-437

Perfect score: 1831

Sequence: 1 MESSGRPSLCQFILLTTSV.....PKKPCICQATRIPLVNS 352

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1831	100.0	352	3 AAB25794	Aab25794 Human sec
2	1831	100.0	352	4 AAE06602	Aae06602 Human pro
3	1831	100.0	352	4 ABB50174	Abb50174 Human tra
4	1831	100.0	352	4 AAB75380	Aab75380 Human sec
5	1831	100.0	352	5 ABP61461	Abp61461 Human NF-
6	1831	100.0	352	5 ABB97404	Abb97404 Novel hum
7	1831	100.0	352	8 ADP19181	Adp19181 Human sec
8	1826	99.7	392	6 AAU15848	Aau15848 Human nov
9	1826	99.7	392	6 ABUS4917	Abu54917 Human nov
10	1824	99.6	352	5 ABP61460	Abp61460 Human NF-
11	1818	99.3	352	4 ABA94353	Aab94353 Human pro
12	1328.5	72.6	263	6 ABJ38698	Abj38698 Human nuc
13	901	49.2	174	4 AAU16309	Aau16309 Human nov
14	901	49.2	174	6 ABUS5378	Abu5378 Human nov
15	854	46.6	165	5 ABB30304	Abb90304 Human pol
16	445	24.3	91	2 AAY13137	Aay13137 Human sec
17	442.5	24.2	338	4 ABB57873	Abb57873 Drosophil
18	439	24.0	90	2 AAY11645	Aay11645 Human 5'
19	437	23.9	88	2 AAY11871	Aay11871 Human 5'
20	201	11.0	383	3 AAG32339	Aag32339 Arabidops
21	200	10.9	383	3 AAG10096	Aag10096 Arabidops
22	200	10.9	399	3 AAG10095	Aag10095 Arabidops
23	188.5	10.3	332	7 ADD25209	Add25209 Fertility
24	188.5	10.3	332	8 ADN61224	Adn61224 Radish nu
25	187.5	10.2	242	3 AAG32340	Aag32340 Arabidops

ALIGNMENTS

RESULT 1									
AAB25794									
ID	AAB25794	standard; protein; 352 AA.							
XX	AC	AAB25794;							
XX	AC	AAB25794;							
DT	28-NOV-2000	(first entry)							
XX	DE	Human secreted protein SEQ ID #106.							
XX	DE	Human secreted protein SEQ ID #106.							
KW	KW	Human; secreted protein; forensic procedure; gene therapy;							
KW	KW	chromosome mapping; cancer; autoimmune disease; cardiovascular disorder;							
KW	KW	cystic fibrosis; hypothyroidism; immunological disorder; amyloidosis;							
KW	KW	brain disorder; skeletal muscle disorder; eye disorder; obesity;							
KW	KW	mitochondriopathy; diabetes; atherosclerosis; Alzheimer's disease;							
KW	KW	neurodegenerative disorder; graft rejection; dementia; hyperlipidaemia;							
XX	XX	septic shock; impotence.							
OS	OS	Homo sapiens.							
XX	XX	WO200037491-A2.							
XX	XX	29-JUN-2000.							
PD	PD	20-DEC-1999; 99WO-IB002058.							
XX	XX	22-DEC-1998; 98US-0113686P.							
XX	XX	25-JUN-1999; 99US-0141032P.							
PA	PA	(GEST) GENSET.							
XX	XX	Bougueleret L, Dumas J, Duclert A;							
PI	PI	WPI; 2000-442637/38.							
XX	XX	N-PSDB; AAA87756.							
DR	DR	Polynucleotides and polypeptides encoding proteins with signal peptides,							
XX	XX	useful in diagnostic, forensic, gene therapy and chromosome mapping							
PT	PT	procedures.							
XX	XX	Claim 9; Page 279-280; 306pp; English.							
PS	PS	This sequence represents a human secreted protein amino acid sequence.							
XX	XX	The invention relates to sequences AAA87725-A87774 which encode human							
CC	CC	secreted proteins AAB25763-B25812. The proteins include signal peptides.							
CC	CC	Included in the invention are a host cell containing one of the cDNA							
CC	CC	sequences, and a purified antibody capable of binding to one of the							
CC	CC	secreted proteins. Also contained in the invention are methods for							

26	186.5	10.2	242	3	AAG10097	Aag10097 Arabidops
27	177	9.7	206	3	AAG32341	Aag32341 Arabidops
28	163	8.9	257	4	AAE10909	Aae10909 Human gen
29	163	8.9	372	3	AAV56882	Aav56882 Human apo
30	155.5	8.5	303	6	ADA54907	Ada54907 Human pro
31	155.5	8.5	696	6	ADA54857	Ada54857 Human pro
32	155.5	8.5	723	3	AAV96308	Aav96308 Human RIN
33	155.5	8.5	723	5	ABG32886	Abg32886 Human RIN
34	155.5	8.5	723	7	ADB64418	Adb64418 Human pro
35	153	8.4	275	3	AAG20337	Aag20337 Arabidops
36	151	8.2	700	4	ABB66133	Abb66133 Drosophil
37	148.5	8.1	115	4	AAU00781	Aau00781 Human apo
38	148.5	8.1	272	7	ADM05923	Adm05923 Human pro
39	145.5	7.9	336	5	ABP61002	Abp61002 Novel hum
40	144.5	7.9	438	2	AAW04583	Aaw04583 Human inh
41	144.5	7.9	618	2	AAW19746	Aaw19746 Human inh
42	144.5	7.9	618	2	AAW13545	Aaw13545 Human c-I
43	144.5	7.9	618	2	AAV33998	Aav33998 Human cel
44	144.5	7.9	618	6	ABP72160	Abp72160 Human inh
45	144.5	7.9	618	7	ADB80949	Adb80949 RING-SH c

CC storing the sequence data on a computer system, and a method for
 CC identifying features of the cDNA sequences using a computer programme.
 CC The cDNAs are useful for expressing secreted proteins or fragments to
 CC obtain antibodies capable of specifically binding to the secreted
 CC proteins. The cDNAs may also be useful in diagnostic, forensic, gene
 CC therapy and chromosome mapping procedures and may be used to design
 CC expression vectors and secretion vectors. The proteins of the invention
 CC may be used to treat diseases including cancer, autoimmune diseases,
 CC cardiovascular disorders, cystic fibrosis, hypothyroidism, immunological
 CC disorders, amyloidosis, brain disorders, skeletal muscle disorders, eye
 CC disorders, obesity, mitochondrialopathies, diabetes, atherosclerosis,
 CC neurodegenerative disorders, graft rejection, Alzheimer's disease,
 CC dementia, hyperlipidaemia, septic shock and impotence
 XX
 SQ Sequence 352 AA;
 Query Match 100.0%; Score 1831; DB 3; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.2e-168;
 Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MESGRPSLCQFILLGTTSSVTAALYSVYRQKARVSQELKGAKVHLGEDLKSILSEAPG 60
 DB 1 MESGRPSLCQFILLGTTSSVTAALYSVYRQKARVSQELKGAKVHLGEDLKSILSEAPG 60
 QY 61 KCVYAVTEGAVRSVKETLNSQFVENCCKGVQRLTLQEHKVMWNTTTLWDCSKIIHOR 120
 DB 61 KCVYAVTEGAVRSVKETLNSQFVENCCKGVQRLTLQEHKVMWNTTTLWDCSKIIHOR 120
 QY 121 TINTVPFDLVPHEGDVDAVRVLKPLDSVDLGLTVEYKFPSPISQFTDVIHYSGERPK 180
 DB 121 TINTVPFDLVPHEGDVDAVRVLKPLDSVDLGLTVEYKFPSPISQFTDVIHYSGERPK 180
 QY 181 GIQTEEMLKVGATLTGVGELVLDNNSVRLQPPKQGMQYILSSQDFSLQRESSVRLW 240
 DB 181 GIQTEEMLKVGATLTGVGELVLDNNSVRLQPPKQGMQYILSSQDFSLQRESSVRLW 240
 QY 241 KVLALVFGFATCATLFFILRKQYLQRELRKQMEBFQEHQAQLLSRAKPEDRESLKS 300
 DB 241 KVLALVFGFATCATLFFILRKQYLQRELRKQMEBFQEHQAQLLSRAKPEDRESLKS 300
 QY 301 ACVWCLSSFKSCVFLQECGHVCSCTCYRALPEPKKPCICRQAITRVIPLYS 352
 DB 301 ACVWCLSSFKSCVFLQECGHVCSCTCYRALPEPKKPCICRQAITRVIPLYS 352
 RESULT 2
 ID AAE06602
 XX AAE06602 standard; protein; 352 AA.
 AC AAE06602;
 XX
 DT 25-SEP-2001 (first entry)
 DE
 XX Human protein having hydrophobic domain, HP10649.
 XX Human; hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 KW contraceptive; antiinfertility; antiinflammatory.
 XX
 OS Homo sapiens.
 XX
 PN WO200149728-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-0P009359.
 XX
 PR 06-JAN-2000; 2000JP-00000585.
 PR 06-JAN-2000; 2000JP-00000588.

PR 11-JAN-2000; 2000JP-00002299.
 PR 03-FEB-2000; 2000JP-00026862.
 PR 03-MAR-2000; 2000JP-00058367.
 XX
 PA (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 XX
 PI Kato S, Kimura T;
 XX WPI; 2001-418355/44.
 DR N-PSDB; AAD12597.
 XX
 PT Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer,
 PT Alzheimer's and inflammation.
 XX
 Claim 1; Page 122; 563pp; English.
 PS
 XX The present sequence is human protein with hydrophobic domain, HP10649.
 CC The polynucleotide and polypeptide of the invention may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate polypeptide expression. The polynucleotides may be used to
 CC produce the polypeptide, by inserting the nucleic acids into a host cell
 CC and culturing the cell to express the protein. The polynucleotides and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays and also used in gene therapy. The polypeptides may also be used
 CC as antigens in the production of antibodies and in assays to identify
 CC modulators of polypeptide expression and activity. The polypeptides and
 CC nucleic acids may be used as nutritional supplements, to modulate
 CC cytokine and cell proliferation activity, to modulate immune stimulation
 CC or suppression (e.g. for the treatment of microbial infections and
 CC autoimmune disorders such as multiple sclerosis, rheumatoid arthritis and
 CC insulin-dependent diabetes), to modulate haematopoiesis, to modulate
 CC tissue growth activity (e.g. for the treatment of Parkinson's disease,
 CC Huntington's disease and Alzheimer's disease), to modulate activin and
 CC inhibin activity (e.g. for controlling fertility), to modulate
 CC chemotactic and chemokinetic activity, to modulate haemostatic and
 CC thrombolytic activity, to modulate receptor ligand activity, to modulate
 CC inflammation and to inhibit tumour growth
 XX
 SQ Sequence 352 AA;
 Query Match 100.0%; Score 1831; DB 4; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.2e-168;
 Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MESGRPSLCQFILLGTTSSVTAALYSVYRQKARVSQELKGAKVHLGEDLKSILSEAPG 60
 DB 1 MESGRPSLCQFILLGTTSSVTAALYSVYRQKARVSQELKGAKVHLGEDLKSILSEAPG 60
 QY 61 KCVYAVTEGAVRSVKETLNSQFVENCCKGVQRLTLQEHKVMWNTTTLWDCSKIIHOR 120
 DB 61 KCVYAVTEGAVRSVKETLNSQFVENCCKGVQRLTLQEHKVMWNTTTLWDCSKIIHOR 120
 QY 121 TINTVPFDLVPHEGDVDAVRVLKPLDSVDLGLTVEYKFPSPISQFTDVIHYSGERPK 180
 DB 121 TINTVPFDLVPHEGDVDAVRVLKPLDSVDLGLTVEYKFPSPISQFTDVIHYSGERPK 180
 QY 181 GIQTEEMLKVGATLTGVGELVLDNNSVRLQPPKQGMQYILSSQDFSLQRESSVRLW 240
 DB 181 GIQTEEMLKVGATLTGVGELVLDNNSVRLQPPKQGMQYILSSQDFSLQRESSVRLW 240
 QY 241 KVLALVFGFATCATLFFILRKQYLQRELRKQMEBFQEHQAQLLSRAKPEDRESLKS 300
 DB 241 KVLALVFGFATCATLFFILRKQYLQRELRKQMEBFQEHQAQLLSRAKPEDRESLKS 300
 QY 301 ACVWCLSSFKSCVFLQECGHVCSCTCYRALPEPKKPCICRQAITRVIPLYS 352
 DB 301 ACVWCLSSFKSCVFLQECGHVCSCTCYRALPEPKKPCICRQAITRVIPLYS 352
 RESULT 3
 ABB50174

ID ABB50174 standard; protein; 352 AA.
 AC ABB50174;
 DT 05-FEB-2002 (first entry)
 DE Human transcription factor TRFX-25.
 DE Human; transcription factor; TRFX; cell proliferative disease;
 KW autoimmune disease; inflammation; neurological disease;
 KW developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;
 KW neuroprotective; antiinflammatory; gene therapy.
 OS Homo sapiens.
 XX WO200172777-A2.
 XX 04-OCT-2001.
 PF 13-MAR-2001; 2001WO-US008117.
 XX 13-MAR-2000; 2000US-0188986P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Hillman JL, Baughn MR, Yue H, Lal P, Lu DAM, Patterson C;
 PI Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;
 PI Reddy R;
 XX WPI: 2001-570896/64.
 DR N-PSDB; ABA82998.
 XX Novel transcription factor polypeptides, used to treat diseases
 PT associated with altered activity and expression of TRFX, and to screen
 PT for agents capable of modulating its activity.
 XX Claim 1; Page 168-169; 327pp; English.
 CC The present sequence is the protein sequence for a human transcription
 CC factor. The transcription factor and its coding sequence are useful in
 CC the diagnosis, treatment and prevention of diseases associated with
 CC altered expression of the transcription factor e.g. cell proliferative,
 CC autoimmune/inflammatory, neurological and developmental disorders. A
 CC number of specific disorders/diseases are given in the specification,
 CC including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,
 CC allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic
 CC dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,
 CC Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,
 CC psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative
 CC colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's
 CC disease, stroke, and viral, bacterial, fungal and protozoal infections
 XX SQ Sequence 352 AA;
 Query Match 100.0%; Score 1831; DB 4; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.2e-168;
 Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MESGGRSLCQFILLGTTTAAALYSVYRQKARVSQELGAKKXVHLGDLKSLSEAPG 60
 DB 1 MESGGRSLCQFILLGTTTAAALYSVYRQKARVSQELGAKKXVHLGDLKSLSEAPG 60
 QY 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIOQLTLOEHKVMWNRTHLWNCDSKIIHQ 120
 DB 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIOQLTLOEHKVMWNRTHLWNCDSKIIHQ 120
 QY 121 TMTVPDLVPHEDGVAVRVLPKLDSDVLGLTVEYKFKHPSIQSFTDVI GHYISGERPK 180
 DB 121 TMTVPDLVPHEDGVAVRVLPKLDSDVLGLTVEYKFKHPSIQSFTDVI GHYISGERPK 180
 QY 181 GIOETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQRESSVRLW 240
 DB 181 GIOETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQRESSVRLW 240
 QY 241 KVALVFGFATCATLFFILRKQYLRQERLRLKQMOEFOEHAQLLSRAKPEDRESLKS 300
 DB 241 KVALVFGFATCATLFFILRKQYLRQERLRLKQMOEFOEHAQLLSRAKPEDRESLKS 300

QY 241 KVALVFGFATCATLFFILRKQYLRQERLRLKQMOEFOEHAQLLSRAKPEDRESLKS 300
 DB 241 KVALVFGFATCATLFFILRKQYLRQERLRLKQMOEFOEHAQLLSRAKPEDRESLKS 300
 QY 301 ACVVCLSSPKSCVFLGCGHVCSTECYRALPEPKKCPICRQAITRVIPLNS 352
 DB 301 ACVVCLSSPKSCVFLGCGHVCSTECYRALPEPKKCPICRQAITRVIPLNS 352
 RESULT 4
 AAB75380
 ID AAB75380 standard; protein; 352 AA.
 XX AAB75380;
 XX 05-APR-2001 (first entry)
 DT Human secreted protein #39.
 DE Human secreted protein #39.
 XX Secreted protein; prevention; treatment; diagnosis; disease; infection.
 OS Homo sapiens.
 XX WO200100806-A2.
 XX 04-JAN-2001.
 PF 21-JUN-2000; 2000WO-IB000951.
 PR 25-JUN-1999; 99US-0141032P.
 PR 21-DEC-1999; 99US-00469099.
 XX (GEST) GENSET.
 PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
 XX WPI: 2001-071487/08.
 CC 49 Secreted proteins and the nucleic acids encoding them, useful in gene
 CC therapy and for detecting similar sequences in samples.
 XX Claim 10; Page 299-300; 307pp; English.
 CC The present invention relates to 49 Secreted proteins and the cDNAs
 CC encoding them. The protein and nucleic acids may be used in the
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate protein expression
 XX SQ Sequence 352 AA;
 Query Match 100.0%; Score 1831; DB 4; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.2e-168;
 Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MESGGRSLCQFILLGTTTAAALYSVYRQKARVSQELGAKKXVHLGDLKSLSEAPG 60
 DB 1 MESGGRSLCQFILLGTTTAAALYSVYRQKARVSQELGAKKXVHLGDLKSLSEAPG 60
 QY 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIOQLTLOEHKVMWNRTHLWNCDSKIIHQ 120
 DB 61 KCPYAVIEGAVRSVKETLNSQFVENCCKGVIOQLTLOEHKVMWNRTHLWNCDSKIIHQ 120
 QY 121 TMTVPDLVPHEDGVAVRVLPKLDSDVLGLTVEYKFKHPSIQSFTDVI GHYISGERPK 180
 DB 121 TMTVPDLVPHEDGVAVRVLPKLDSDVLGLTVEYKFKHPSIQSFTDVI GHYISGERPK 180
 QY 181 GIOETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQRESSVRLW 240
 DB 181 GIOETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQRESSVRLW 240
 QY 241 KVALVFGFATCATLFFILRKQYLRQERLRLKQMOEFOEHAQLLSRAKPEDRESLKS 300
 DB 241 KVALVFGFATCATLFFILRKQYLRQERLRLKQMOEFOEHAQLLSRAKPEDRESLKS 300

Db 241 KVLALVFGFATCATLFFILRKQYLQORERLRLKQMQEFOHEAQLLSRAKPEDRESLKS 300
ABP61461
Qy 301 ACVVCLSSFKSCVFLECGHVCSTECYRALPEPKKPCICROAIRVPIPLNS 352
Db 301 ACVVCLSSFKSCVFLECGHVCSTECYRALPEPKKPCICROAIRVPIPLNS 352
RESULT 5
ID ABP61461 standard; protein; 352 AA.
XX ABP61461;
XX 30-SEP-2002 (first entry)
XX Human NF-KB activating protein SEQ ID NO 75.
XX Human; NF-kB; nuclear factor kappa B; mouse; antiinflammatory;
KW immunomodulator; cytostatic; antiinfective; osteopathic; nootropic;
KW neuroprotective; anti-HIV; autoimmune disease; cancer; infection;
KW bone disease; AIDS; neurodegenerative disease; ischaemic disorder.
XX Homo sapiens.
XX WO200253737-A1.
XX 11-JUL-2002.
XX 25-DEC-2001; 2001WO-JP011389.
XX 28-DEC-2000; 2000JP-00402288.
XX 26-MAR-2001; 2001JP-00089912.
XX 24-AUG-2001; 2001JP-00254018.
XX (ASAH) ASAH KASEI KOGYO KK.
XX Matsuda A, Honda G, Muramatsu S, Nagano Y;
XX WPI; 2002-583617/62.
XX N-PSDB; ABQ91949.
XX NF-approximatelykB activating gene and expressed protein, applicable in
PT diagnosis and screening inhibitors or promoters to control excessive
PT activation or inhibition for treating e.g. inflammations, autoimmune
PT diseases and cancer.
XX Claim 1; Page 349-352; 841pp; Japanese.
XX The invention relates to a purified protein (I), comprising one of 90
CC fully defined sequences (ABP61424-ABP61513) or a protein based on any of
CC the sequences but with some amino acids deleted, substituted or added and
CC with a NF-kB (nuclear factor kappa B) activating effect. The protein and
CC encoding gene (ABQ91912-ABQ92001) are useful in diagnosis and screening
CC inhibitors or promoters to control excessive activation or inhibition and
CC for treating e.g. inflammations, autoimmune diseases, cancers,
CC infections, bone diseases, AIDS, neurodegenerative diseases or ischaemic
CC disorders
XX Sequence 352 AA;
Query Match 100.0%; Score 1831; DB 5; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.2e-168;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MESGGRSLCQFILLGTTTSVVTALYSVYRQKARVSQELKGAKKXVHLGEDLKSILSPAG 60
Db 1 MESGGRSLCQFILLGTTTSVVTALYSVYRQKARVSQELKGAKKXVHLGEDLKSILSPAG 60
Qy 61 KCVPYAVIEGAVRSVKETLNSQFVENCXGVIQRLTLOEHKQVWNRTHLWNCCKI1HOR 120
Db 61 KCVPYAVIEGAVRSVKETLNSQFVENCXGVIQRLTLOEHKQVWNRTHLWNCCKI1HOR 120
Qy 121 TNTVPFDLVPHEDGVDVAVRVLKPLDSVDLGLTVEYKFPESIQSFTDVIHGYISGRPK 180

Db 121 TNTVPFDLVPHEDGVDVAVRVLKPLDSVDLGLTVEYKFPESIQSFTDVIHGYISGRPK 180
Qy 181 GIQTEEMLKVGATLTGVELVLDNNSVRLOPPKQGMQYLLSSODFSLQROESSVRLW 240
Db 181 GIQTEEMLKVGATLTGVELVLDNNSVRLOPPKQGMQYLLSSODFSLQROESSVRLW 240
Qy 241 KVLALVFGFATCATLFFILRKQYLQORERLRLKQMQEFOHEAQLLSRAKPEDRESLKS 300
Db 241 KVLALVFGFATCATLFFILRKQYLQORERLRLKQMQEFOHEAQLLSRAKPEDRESLKS 300
Qy 301 ACVVCLSSFKSCVFLECGHVCSTECYRALPEPKKPCICROAIRVPIPLNS 352
Db 301 ACVVCLSSFKSCVFLECGHVCSTECYRALPEPKKPCICROAIRVPIPLNS 352
RESULT 6
ABB97404
ID ABB97404 standard; protein; 352 AA.
XX ABB97404;
XX 27-JUN-2002 (first entry)
XX Novel human protein SEQ ID NO: 672.
XX Human; antianaemic; vulnary; antiinflammatory; immunomodulator;
KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag.
XX Homo sapiens.
XX WO200222660-A2.
XX 21-MAR-2002.
XX 10-SEP-2001; 2001WO-US026015.
XX 11-SEP-2000; 2000US-00659671.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-292408/33.
XX N-PSDB; ABN32590.
XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.
XX Claim 20; SEQ ID NO 672; 509pp; English.
XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a protein of the invention
XX Sequence 352 AA;
Query Match 100.0%; Score 1831; DB 5; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.2e-168;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MESGGRSLCQFILLGTTTSVVTALYSVYRQKARVSQELKGAKKXVHLGEDLKSILSPAG 60

Db 1 MESGRPSLCQFILLGTTSVVTAALYSVYRKARVSQELKGAKVHGLGDLKSLSEAPG 60
QY 61 KCVPAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKIIHOR 120
Db 61 KCVPAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKIIHOR 120
QY 121 TMTVPFDLVPHEGVDVAVRVLKPLDSVDLGLTGVYVEKFPSTQSFDTVIGHYISGERPK 180
Db 121 TMTVPFDLVPHEGVDVAVRVLKPLDSVDLGLTGVYVEKFPSTQSFDTVIGHYISGERPK 180
QY 181 GQETEEMLVKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDFDLSLLQRESSVRLW 240
Db 181 GQETEEMLVKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDFDLSLLQRESSVRLW 240
QY 241 KVLALVFGFATCATLFFILRKQYLQORERLRLKQOEERFQHEAQLLSRAKPEDRESLKS 300
Db 241 KVLALVFGFATCATLFFILRKQYLQORERLRLKQOEERFQHEAQLLSRAKPEDRESLKS 300
QY 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352
Db 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352

RESULT 7

ADP19181
ID ADP19181 standard; protein; 352 AA.

AC ADP19181;
XX

DT 26-AUG-2004 (first entry)
DE Human secreted polypeptide #32.

XX Human; secreted protein; genetic disease.
XX Homo sapiens.

FN US2004110939-A1.
XX 10-JUN-2004.

XX 15-OCT-2001; 2001US-00978360.
XX 17-DEC-1998; 98WO-IB002122.

XX 09-FEB-1999; 99WO-IB000282.
XX 21-JUN-2000; 2000WO-IB000951.

XX 15-SEP-2000; 2000US-00663600.
XX (GIST) GENSET SA.

XX Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;
XX Duclert A;

XX WPI; 2004-440404/41.
XX N-PSDB; ADP18776.

PT New isolated polynucleotide encoding secreted polypeptide, useful for
PT gene therapy, or in diagnostic procedures to identify individuals having
PT genetic diseases resulting from abnormal expression of the genes.
XX Claim 2; SEQ ID NO 437; 113pp; English.

XX The invention relates to human cDNA sequences that encode human secreted
CC proteins. The invention also relates to an antibody that specifically
CC binds to a polypeptide of the invention and a method of binding the
CC polypeptide to an antibody. The polynucleotides are useful for expressing
CC the entire secreted proteins which they encode and for distinguishing
CC human tissues and cells from non-human tissues and cells, and for
CC distinguishing between human tissues and cells that do or do not express
CC the polynucleotides comprising the cDNAs. The polynucleotides and
CC polypeptides are useful in forensic procedures or diagnostic procedures
CC to identify individuals with genetic diseases resulting from abnormal
CC expression of the genes corresponding to the cDNAs. The sequences are

CC also useful in gene therapy to control or treat genetic diseases. This
CC sequence represents a human secreted polypeptide of the invention. Note:
CC the sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html.

XX Sequence 352 AA;

Query Match 100.0%; Score 1831; DB 8; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.2e-168;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESGRPSLCQFILLGTTSVVTAALYSVYRKARVSQELKGAKVHGLGDLKSLSEAPG 60
Db 1 MESGRPSLCQFILLGTTSVVTAALYSVYRKARVSQELKGAKVHGLGDLKSLSEAPG 60

QY 61 KCVPAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKIIHOR 120
Db 61 KCVPAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMWNRTHLWDCSKIIHOR 120

QY 121 TMTVPFDLVPHEGVDVAVRVLKPLDSVDLGLTGVYVEKFPSTQSFDTVIGHYISGERPK 180
Db 121 TMTVPFDLVPHEGVDVAVRVLKPLDSVDLGLTGVYVEKFPSTQSFDTVIGHYISGERPK 180

QY 181 GQETEEMLVKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDFDLSLLQRESSVRLW 240
Db 181 GQETEEMLVKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDFDLSLLQRESSVRLW 240

QY 241 KVLALVFGFATCATLFFILRKQYLQORERLRLKQOEERFQHEAQLLSRAKPEDRESLKS 300
Db 241 KVLALVFGFATCATLFFILRKQYLQORERLRLKQOEERFQHEAQLLSRAKPEDRESLKS 300

QY 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352
Db 301 ACVCLSSFKSCVFLGCHVCSCTCYRALPEPKKPCICROAITRVIPLVNS 352

RESULT 8

AAU15848
ID AAU15848 standard; protein; 392 AA.

XX AC AAU15848;
XX

XX 07-NOV-2001 (first entry)
XX

DE Human novel secreted protein, Seq ID 801.
XX

XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.

XX Homo sapiens.
OS

XX WO200155322-A2.
XX

XX 02-AUG-2001.
XX

XX 17-JAN-2001; 2001WO-US001341.
XX

XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234597P.
 PR 25-SEP-2000; 2000US-0234598P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-488783/53.
 DR N-PSDB; AAS25835.
 XX
 XX New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 XX
 XX Claim 11; SEQ ID NO 801; 980pp; English.
 PS
 XX The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used

CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 99.7%; Score 1826; DB 4; Length 392;
Best Local Similarity 99.7%; Pred. No. 4.2e-168;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESSGRPSLCOFILLGTTSSVTAALYSVYRKARVSQELKGAKVHLGDLKSLSEAPG 60
DB 41 MESSGRPSLCOFILLGTTSSVTAALYSVYRKARVSQELKGAKVHLGDLKSLSEAPG 100
QY 61 KCVPVAVIEGAVRSVKETLNSQFVENCKGVQLRTLQEHKVMWNTTHLWDCSKIIHQ 120
DB 101 KCVPVAVIEGAVRSVKETLNSQFVENCKGVQLRTLQEHKVMWNTTHLWDCSKIIHQ 160
QY 121 TNPVFDLVPHEHDGVAVRVLPKLDSDVLGLETVYKPHSIQSFDTVIGHYISGERPK 180
DB 161 TNPVFDLVPHEHDGVAVRVLPKLDSDVLGLETVYKPHSIQSFDTVIGHYISGERPK 220
QY 181 GIOETEEMKVGATLGTGVELVDNNSVRLQPKQMGQYVYSSQDFDLSLQRESSVRLW 240
DB 221 GIOETEEMKVGATLGTGVELVDNNSVRLQPKQMGQYVYSSQDFDLSLQRESSVRLW 280
QY 241 KVLALVFGFATCATLFFILRKQYLRQLRKLKQMEFFQEHQAQLLSRAKPEDRESLKS 300
DB 281 KVLALVFGFATCATLFFILRKQYLRQLRKLKQMEFFQEHQAQLLSRAKPEDRESLKS 340
QY 301 ACVCLSSFKSCVPLECGHVCSTCYRALPEPKKPCICROAITRVIPLVNS 352
DB 341 ACVCLSSFKSCVPLECGHVCSTCYRALPEPKKPCICROAITRVIPLVNS 392

RESULT 9
ABUS4917
ID ABUS4917 standard; protein; 392 AA.

AC ABUS4917;

DT 18-MAR-2003 (first entry)

XX Human novel polypeptide #4.

XX Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.

XX Homo sapiens.

OS US2002132753-A1.

XX 19-SEP-2002..

XX 17-JAN-2001; 2001US-00764864.
PF 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX Rosen CA, Ruben SM, Barash SC;
PI MPI, 2003-147444/14.
XX N-PSDB; ABX73176.
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX Claim 11; SEQ ID NO 801; 402pp; English.
PS The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune

CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
 CC ABU55748 represent human novel polypeptides of the invention

XX SQ Sequence 392 AA;

Query Match 99.7%; Score 1826; DB 6; Length 392;
 Best Local Similarity 99.7%; Pred. No. 4.2e-168;
 Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MESGGRPSCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKVHLGDLKSLSEAPG 60
 DB 41 MESGGRPSCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKVHLGDLKSLSEAPG 100
 QY 61 KCVYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMVNRTHLWDCSKIIHOR 120
 DB 101 KCVYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMVNRTHLWDCSKIIHOR 160
 QY 121 TMTVPFDLVPHEDGVDVAVRVLKPLDSVDLGLTETVYKPHPSIQSFDTVIGHYISGERPK 180
 DB 161 TMTVPFDLVPHEDGVDVAVRVLKPLDSVDLGLTETVYKPHPSIQSFDTVIGHYISGERPK 220
 QY 181 GQETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQRESSVRLW 240
 DB 221 GQETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQRESSVRLW 280
 QY 241 KVLALVFGFATCATLFFILRKQYLQERLRKQMQEFPQEHQAQLLSRAKPEDRESLKS 300
 DB 281 KVLALVFGFATCATLFFILRKQYLQERLRKQMQEFPQEHQAQLLSRAKPEDRESLKS 340
 QY 301 ACVCLSSPKSCVFLECGHVCSTECYRALPPEPKKPCICROAITRVIPLYS 352
 DB 341 ACVCLSSPKSCVFLECGHVCSTECYRALPPEPKKPCICROAITRVIPLYS 392

RESULT 10
 ID ABP61460 standard; protein; 352 AA.

XX AC ABP61460;
 XX DT 30-SEP-2002 (first entry)
 XX DE Human NF-kB activating protein SEQ ID NO 73.

XX Human; NF-kB; nuclear factor kappa B; mouse; antiinflammatory;
 KW immunomodulator; cytostatic; antiinfective; osteopathic; nootropic;
 KW neuroprotective; anti-HIV; autoimmune disease; cancer; infection;
 KW bone disease; AIDS; neurodegenerative disease; ischaemic disorder.

XX OS Homo sapiens.
 XX PN WO200253737-A1.
 XX PD 11-JUL-2002.
 XX PF 25-DEC-2001; 2001WO-JF011389.
 XX PR 28-DEC-2000; 2000JP-00402288.
 XX PR 26-MAR-2001; 2001JP-00089912.
 XX PR 24-AUG-2001; 2001JP-00254018.
 XX PA (ASAH) ASahi Kasei Kogyo KK.

PI Matsuda A, Honda G, Muramatsu S, Nagano Y;
 XX WPI; 2002-583617/62.
 DR N-P8DB; ABQ91948.
 XX NF-approximatelykB activating gene and expressed protein, applicable in
 PT diagnosis and screening inhibitors or promoters to control excessive
 PT activation or inhibition for treating e.g. inflammations, autoimmune
 PT diseases and cancer.

XX Claim 1; Page 342-344; 841pp; Japanese.

XX The invention relates to a purified protein (I), comprising one of 90
 CC fully defined sequences (ABP61424-ABP61513) or a protein based on any of
 CC the sequences but with some amino acids deleted, substituted or added and
 CC with a NF-kB (nuclear factor kappa B) activating effect. The protein and
 CC encoding gene (ABQ91912-ABQ92001) are useful in diagnosis and screening
 CC inhibitors or promoters to control excessive activation or inhibition and
 CC for treating e.g. inflammations, autoimmune diseases, cancers,
 CC infections, bone diseases, AIDS, neurodegenerative diseases or ischaemic
 CC disorders

XX SQ Sequence 352 AA;

Query Match 99.6%; Score 1824; DB 5; Length 352;
 Best Local Similarity 99.7%; Pred. No. 5.6e-168;
 Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MESGGRPSCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKVHLGDLKSLSEAPG 60
 DB 1 MESGGRPSCQFILLGTTTSVVTAAALYSVYRQKARVSQELKGAKVHLGDLKSLSEAPG 60
 QY 61 KCVYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMVNRTHLWDCSKIIHOR 120
 DB 61 KCVYAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKVMVNRTHLWDCSKIIHOR 120
 QY 121 TMTVPFDLVPHEDGVDVAVRVLKPLDSVDLGLTETVYKPHPSIQSFDTVIGHYISGERPK 180
 DB 121 TMTVPFDLVPHEDGVDVAVRVLKPLDSVDLGLTETVYKPHPSIQSFDTVIGHYISGERPK 180
 QY 181 GQETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQRESSVRLW 240
 DB 181 GQETEMLKVGATLTGVELVLDNNSVRLQPPKQGMQYLLSQDPSLLQRESSVRLW 240
 QY 241 KVLALVFGFATCATLFFILRKQYLQERLRKQMQEFPQEHQAQLLSRAKPEDRESLKS 300
 DB 241 KVLALVFGFATCATLFFILRKQYLQERLRKQMQEFPQEHQAQLLSRAKPEDRESLKS 300
 QY 301 ACVCLSSPKSCVFLECGHVCSTECYRALPPEPKKPCICROAITRVIPLYS 352
 DB 301 ACVCLSSPKSCVFLECGHVCSTECYRALPPEPKKPCICROAITRVIPLYS 352

RESULT 11
 AAB94353
 ID AAB94353 standard; protein; 352 AA.
 XX AC AAB94353;
 XX DT 26-JUN-2001 (first entry)
 XX DE Human protein sequence SEQ ID NO:14873.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-00116126.
 XX

```
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118772.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 14873; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dr primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95993
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 352 AA;
XX
XX Query Match 99.3%; Score 1818; DB 4; Length 352;
XX Best Local Similarity 99.7%; Pred. No. 2.1e-167;
XX Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MESGRPSLCQFILLGTTTSVVTAAALYSYVYQKARVSELKGAKKVHLGEDLKSILSEAPG 60
XX 1 MESGRPSLCQFILLGTTTSVVTAAALYSYVYQKARVSELKGAKKVHLGEDLKSILSEAPG 60
XX
XX 61 KCVPAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNTTTLWNCDSKIIHQH 120
XX 61 KCVPAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNTTTLWNCDSKIIHQH 120
XX
XX 121 TNTVPFDLVPHEDGVDVAVRLKPLDSVDLGLTVEYKFPSPISQFTDVIHYISGERPK 180
XX 121 TNTVPFDLVPHEDGVDVAVRLKPLDSVDLGLTVEYKFPSPISQFTDVIHYISGERPK 180
XX
XX 181 GIQETEMMLKVCATLTGVELVDNNSVRLQPPKQGMQYLLSSQDFSLQROESSVRLW 240
XX 181 GIQETEMMLKVCATLTGVELVDNNSVRLQPPKQGMQYLLSSQDFSLQROESSVRLW 240
XX
XX 241 KYLALVFGFATCATLFFILRKQYLOQRERLQKQWERFOEHAQLLSRAKPEDRESLKS 300
XX 241 KYLALVFGFATCATLFFILRKQYLOQRERLQKQWERFOEHAQLLSRAKPEDRESLKS 300
XX
XX 301 ACWCGLSSPKSCVFLKCGHVCCTECYRALPFPKPCICRQAITRVIPLYS 352
XX 301 ACWCGLSSPKSCVFLKCGHVCCTECYRALPFPKPCICRQAITRVIPLYS 352
XX
```

```
RESULT 12
ABU38698
ID ABU38698 standard; protein; 263 AA.
XX
XX AC ABU38698;
XX
XX DT 04-AUG-2003 (first entry)
XX
XX DE Human nucleic acid-associated protein (NAAP) #28.
XX
XX KW Human; gene therapy; nucleic acid-associated protein; NAAP;
KW cell proliferative disorder; cancer; neurological disorder; epilepsy;
KW immune disorder; inflammatory disorder; AIDS; allergy;
KW developmental disorder; Cushing's syndrome.
XX
XX OS Homo sapiens.
XX
XX PN WO2003010329-A2.
XX
XX PD 06-FEB-2003.
XX
XX PF 25-JUL-2002; 2002WO-US023756.
XX
XX PR 26-JUL-2001; 2001US-0308189P.
XX PR 27-JUL-2001; 2001US-0308171P.
XX PR 02-AUG-2001; 2001US-0310139P.
XX PR 03-AUG-2001; 2001US-0309974P.
XX PR 08-AUG-2001; 2001US-0311072P.
XX PR 10-AUG-2001; 2001US-0311642P.
XX PR 10-AUG-2001; 2001US-0311717P.
XX PR 12-OCT-2001; 2001US-0329688P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Tang YT, Nguyen DB, Yao MG, Warren BA, Griffin JA, Ison CH;
XX Forsythe IJ, Becha SD, Yue H, Emerling BM, Wallia NK, Richardson TW;
XX Lee EA, Ramkumar J, Elliott VS, He A, Li JX, Hafalia AJA, Yang J;
XX Sanjanwala MM, Xu Y, Arvizu CS, Gandhi AR, Borowsky ML, Tran UK;
XX Burford N, Sprague WW, Baughn MR, Swarnakar A, Policky JL, Lee SY;
XX Jiang X, Jackson A, Chang H;
XX
XX WPI; 2003-248084/24.
XX DR N-PSDB; ABT42547.
XX
XX New human nucleic acid associated proteins (NAAP), useful for diagnosing,
XX treating and preventing diseases or conditions associated with the
XX aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or
XX infections.
XX
XX Claim 1; Page 229; 263pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
XX nucleic acid-associated proteins (NAAP). The DNA and protein sequences of
XX the invention are useful for diagnosing, treating and preventing diseases
XX or conditions associated with the decreased expression or overexpression
XX of NAAP, such as: cell proliferative disorders (e.g. cancer);
XX neurological disorders (e.g. epilepsy); immune/inflammatory disorders
XX (e.g. AIDS and allergies); and developmental disorders (e.g. Cushing's
XX syndrome). The present amino acid sequence represents a human nucleic
XX acid-associated protein of the invention
XX
XX Sequence 263 AA;
XX
XX Query Match 72.6%; Score 1328.5; DB 6; Length 263;
XX Best Local Similarity 74.7%; Pred. No. 4.5e-120;
XX Matches 263; Conservative 0; Mismatches 0; Indels 89; Gaps 1;
XX
XX 1 MESGRPSLCQFILLGTTTSVVTAAALYSYVYQKARVSELKGAKKVHLGEDLKSILSEAPG 60
XX 1 MESGRPSLCQFILLGTTTSVVTAAALYSYVYQKARVSELKGAKKVHLGEDLKSILSEAPG 60
XX
XX 61 KCVPAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNTTTLWNCDSKIIHQH 120
XX 61 KCVPAVIEGAVRSVKETLNSQFVENCCKGVIQRLTLQEHKMWNTTTLWNCDSKIIHQH 120
XX
```

Db	17									PR	14-AUG-2000;	2000US-0225270P.
										PR	14-AUG-2000;	2000US-0225447P.
										PR	14-AUG-2000;	2000US-0225757P.
Qy	121	TNTVPDLPHEDGVDVAVRLKPLDSVDLGLTETVEKEPHPSIQSTDVIGHYISGERPK	180							PR	14-AUG-2000;	2000US-0225758P.
										PR	14-AUG-2000;	2000US-0225759P.
Db	32	TNTVPDLPHEDGVDVAVRLKPLDSVDLGLTETVEKEPHPSIQSTDVIGHYISGERPK	91							PR	18-AUG-2000;	2000US-0226279P.
										PR	22-AUG-2000;	2000US-0226681P.
Qy	181	GQETHEMLKVGATLTGVGELVDNNSVRLQPPKQGMQYVLSQDFSLQROESSVRLW	240							PR	22-AUG-2000;	2000US-0226868P.
										PR	22-AUG-2000;	2000US-0227182P.
Db	92	GQETHEMLKVGATLTGVGELVDNNSVRLQPPKQGMQYVLSQDFSLQROESSVRLW	151							PR	23-AUG-2000;	2000US-0227009P.
										PR	30-AUG-2000;	2000US-0228924P.
Qy	241	KVALVFGFATCATLFFILRKQYLRQERLRKQMOEFQEHEAQLLSRAKPEDRESLKS	300							PR	01-SEP-2000;	2000US-0229287P.
										PR	01-SEP-2000;	2000US-0229343P.
Db	152	KVALVFGFATCATLFFILRKQYLRQERLRKQMOEFQEHEAQLLSRAKPEDRESLKS	211							PR	01-SEP-2000;	2000US-0229344P.
										PR	01-SEP-2000;	2000US-0229345P.
Qy	301	ACVCLSSFKSCVFLECGHVCSTECYRALPEPKKPCICQAITRVIPLYS	352							PR	05-SEP-2000;	2000US-0229509P.
										PR	05-SEP-2000;	2000US-0229513P.
Db	212	ACVCLSSFKSCVFLECGHVCSTECYRALPEPKKPCICQAITRVIPLYS	263							PR	06-SEP-2000;	2000US-0230437P.
										PR	06-SEP-2000;	2000US-0230438P.
										PR	08-SEP-2000;	2000US-0231242P.
										PR	08-SEP-2000;	2000US-0231243P.
										PR	08-SEP-2000;	2000US-0231244P.
										PR	08-SEP-2000;	2000US-0231413P.
										PR	08-SEP-2000;	2000US-0231414P.
										PR	08-SEP-2000;	2000US-0232080P.
										PR	08-SEP-2000;	2000US-0232081P.
										PR	12-SEP-2000;	2000US-0231968P.
										PR	14-SEP-2000;	2000US-0232397P.
										PR	14-SEP-2000;	2000US-0232398P.
										PR	14-SEP-2000;	2000US-0232399P.
										PR	14-SEP-2000;	2000US-0232400P.
										PR	14-SEP-2000;	2000US-0232401P.
										PR	14-SEP-2000;	2000US-0233063P.
										PR	14-SEP-2000;	2000US-0233064P.
										PR	14-SEP-2000;	2000US-0233065P.
										PR	21-SEP-2000;	2000US-0234223P.
										PR	21-SEP-2000;	2000US-0234274P.
										PR	25-SEP-2000;	2000US-0234997P.
										PR	25-SEP-2000;	2000US-0234998P.
										PR	26-SEP-2000;	2000US-0235484P.
										PR	27-SEP-2000;	2000US-0235834P.
										PR	27-SEP-2000;	2000US-0235836P.
										PR	29-SEP-2000;	2000US-0236327P.
										PR	29-SEP-2000;	2000US-0236367P.
										PR	29-SEP-2000;	2000US-0236368P.
										PR	29-SEP-2000;	2000US-0236369P.
										PR	29-SEP-2000;	2000US-0236370P.
										PR	02-OCT-2000;	2000US-0236802P.
										PR	02-OCT-2000;	2000US-0237037P.
										PR	02-OCT-2000;	2000US-0237038P.
										PR	02-OCT-2000;	2000US-0237039P.
										PR	02-OCT-2000;	2000US-0237040P.
										PR	13-OCT-2000;	2000US-0239935P.
										PR	13-OCT-2000;	2000US-0239937P.
										PR	20-OCT-2000;	2000US-0240960P.
										PR	20-OCT-2000;	2000US-0241221P.
										PR	20-OCT-2000;	2000US-0241785P.
										PR	20-OCT-2000;	2000US-0241786P.
										PR	20-OCT-2000;	2000US-0241787P.
										PR	20-OCT-2000;	2000US-0241808P.
										PR	20-OCT-2000;	2000US-0241809P.
										PR	20-OCT-2000;	2000US-0241826P.
										PR	01-NOV-2000;	2000US-0244617P.
										PR	08-NOV-2000;	2000US-0246474P.
										PR	08-NOV-2000;	2000US-0246475P.
										PR	08-NOV-2000;	2000US-0246476P.
										PR	08-NOV-2000;	2000US-0246477P.
										PR	08-NOV-2000;	2000US-0246478P.
										PR	08-NOV-2000;	2000US-0246523P.
										PR	08-NOV-2000;	2000US-0246524P.
										PR	08-NOV-2000;	2000US-0246525P.
										PR	08-NOV-2000;	2000US-0246526P.
										PR	08-NOV-2000;	2000US-0246527P.

PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244517P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
XX N-PSDB; ABX73637.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 11; SEQ ID NO 1262; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
XX therapy for treating, inhibiting or preventing neural disorders, immune
XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
XX (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
XX heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
XX renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
XX leukemia), inflammatory diseases (e.g. septic shock, bursitis and
XX appendicitis), allergic reactions and conditions (e.g. asthma), blood
XX related disorders (e.g. thrombosis, atherosclerosis and myocardial
XX infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
XX ABU55748 represent human novel polypeptides of the invention
XX
XX Sequence 174 AA;
SQ

Query Match 49.2%; Score 901; DB 6; Length 174;
Best Local Similarity 99.4%; Pred. No. 7.5e-79;
Matches 173; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 179 PKGIQTEEMLVKVGATLTGVGELVLDNNSVRLPPKQGMQYLLSSQDFDLSLLQRESSVR 238
DB 1 PKGIQTEEMLVKVGATLTGVGELVLDNNSVRLPPKQGMQYLLSSQDFDLSLLQRESSVR 60
QY 239 LKVKLVAVFGFATCATLFFILRKQYLQORERLKKQMOEFQHEAQLLSRAKPEDRESL 298
DB 61 LKVKLVAVFGFATCATLFFILRKQYLQORERLKKQMOEFQHEAQLLSRAKPEDRESL 120
QY 299 KSAACVVLSSPKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVPLVNS 352
|||||

DB 121 KSAACVVLSSPKSCVFLGCHVCSCTCYRALPEPKKPCICRQAITRVPLVNS 174
RESULT 15
ABB90304
ID ABB90304 standard; protein; 165 AA.
XX
XX ABB90304;
AC
XX
XX 24-MAY-2002 (first entry)
DT
XX
XX Human polypeptide SEQ ID NO 2680.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antitallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
XX vulvar; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
OS
XX Homo sapiens.
XX
XX WO200190304-A2.
PN
XX
XX 29-NOV-2001.
PD
XX
XX 18-MAY-2001; 2001WO-US016450.
PF
XX
XX 19-MAY-2000; 2000US-0205515P.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Birse CE, Rosen CA;
PI
XX
XX WPI; 2002-122018/16.
DR N-PSDB; ABL90713.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
XX Claim 11; SEQ ID NO 2680; 2081pp + Sequence Listing; English.
PS
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (ABB9040-ABB9044) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemia; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: the sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ

Query Match 46.6%; Score 854; DB 5; Length 165;
Best Local Similarity 99.4%; Pred. No. 2.5e-74;
Matches 164; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 MLKVGATLTGVGELVLDNNSVRLPPKQGMQYLLSSQDFDLSLLQRESSVRLMKVLALVF 247
DB 1 MLKVGATLTGVGELVLDNNSVRLPPKQGMQYLLSSQDFDLSLLQRESSVRLMKVLALVF 60
QY 248 GFATCATLFFILRKQYLQORERLKKQMOEFQHEAQLLSRAKPEDRESLKKACVVCLS 307
|||||

Db 61 GFATCATLFFILRKQVLRQERLRLKOMQBEFOEHEAQLLSRAKPEDXESLKSACVVCLS 120
QY 308 SFKSCVFLECGHVCSTECYRALPEPKKCPICQAITRVIPLYS 352
Db 121 SFKSCVFLECGHVCSTECYRALPEPKKCPICQAITRVIPLYS 165

Search completed: November 6, 2004, 17:38:18
Job time : 160 secs

Db 241 KVLALVFGFATCATLFFILRKQYLQORRLKQKQEEFQHEAQLLRAPKPEDRESLS 300
QY 301 ACVCLSFKSCVFLECGHVCSTCYRALPEPKKPCICRQAITRVIPLYS 352
Db 301 ACVCLSFKSCVFLECGHVCSTCYRALPEPKKPCICRQAITRVIPLYS 352

RESULT 2
US-08-905-223-421
; Sequence 421, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 421:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -23...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq QFILLGTTSVVTA/AL
US-08-905-223-421

Query Match 24.3%; Score 445; DB 3; Length 91;
Best Local Similarity 98.9%; Pred. No. 5.2e-41;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MESGGRSLCQFILLGTTSVVTAALYSVYRQKARVSQELKGAKKHVHGEDLKSILSAPG 60
Db 1 MESGGRSLCQFILLGTTSVVTAALYSVYRQKARVSQELKGAKKHVHGEDLKSILSAPG 60
QY 61 KCVPYAVIEGAVRSVKETLNSQFVENCXGV 90
Db 61 KCVPYAVIEGAVRSVKETLNSQFVENCXGV 90
RESULT 3
US-09-324-455-2

; Sequence 2, Application US/09324455
; Patent No. 6326481
; GENERAL INFORMATION:
; APPLICANT: Iowe, David
; TITLE OF INVENTION: NOVEL MOLECULES OF THE AIP-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/324,455
; FILING DATE: 02-JUN-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/087,761
; FILING DATE: 02-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/069001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-324-455-2

Query Match 8.9%; Score 163; DB 3; Length 372;
Best Local Similarity 30.1%; Pred. No. 5.2e-09;
Matches 40; Conservative 29; Mismatches 48; Indels 16; Gaps 5;
QY 222 SSQDFSLILQRESSVRLMKVLALVFGFATCATLFFILRKQYLQORRLKQKQEEF 279
Db 254 SLSDLSLDDVEGMSVRQLKEILARNFVNYSCCERWELVEK-----VNRLYKEN 303
QY 280 QEHEAQLLRAPKPEDRESLSKACVCLSFKSCVFLECGHVCSTCYRALPEPKKPCIC 339
Db 304 EENQKSYGERLQLODEED-DSLCRICMDAVIDCVLLECGHVMVCTCKGRMSE---CPIC 359
QY 340 RQAITRVIPLYS 352
Db 360 RQYVWRVAVHVKFS 372

RESULT 4
US-09-434-408-2
; Sequence 2, Application US/09434408
; Patent No. 6440697
; GENERAL INFORMATION:
; APPLICANT: Venezia, Domenick
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: RING FINGER PROTEIN ZAPOF3
; FILE REFERENCE: 98-41
; CURRENT APPLICATION NUMBER: US/09/434,408
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: US 60/108,258
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0

Query Match

Db 332 TILVKGNAANIFKNCLKEIDSTLYKNLFVDMKMKYIPTEDYSGLSLEEQRLRLOEE--- 388
QY 282 HEAQLLSRAKPEDRESLSKACVCLSSFKSCVFLGCHVSCCTCYRALPEPKKPCICRQ 341
Db 389 -----RTCKVMDKEVSVVFIPOGHLVWCQEC---APSLRKCPCICRG 427
QY 342 AITRVIPLVNS 352
Db 428 IIKGTVRTFLS 438

RESULT 8
US-08-569-749-2
; Sequence 2, Application US/08569749
; Patent No. 6187557
; GENERAL INFORMATION:
; APPLICANT: Rothe, Mike
; APPLICANT: Goeddel, David V
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/569,749
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brenner, David J.
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: A-62464/DJB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)781-1989
; TELEFAX: (415)398-3249
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-569-749-2

Query Match 7.9%; Score 144.5; DB 3; Length 618;
Best Local Similarity 21.9%; Pred. No. 1.3e-06;
Matches 68; Conservative 50; Mismatches 118; Indels 75; Gaps 13;

QY 82 QFVENCXGVIQRLTLOEHKVMNRTHLWDCSKIIHQRTNTPFDLPVPHEDGVDVAVRV 141
Db 343 EFVDEIQGRYPHLL--EQLLSTSDTTGEENADPPIH-----FGPGESSEDVAVM 391
QY 142 LKPL--DSVDLGL-----ETVVEKHPSTQSF---TDVIGHVISGERPKGIOE---TE 186
Db 392 NTPVKSALGEMGNRDLVKQTQVSKILTTGENTYKTVNDIVSALLNADEKREKEKEQAE 451
QY 187 EMLKVGATLTGVGEL-----VLDN-----NSVRLQPP---KQGMQYVLSQD--FD 227
Db 452 EMASDDLSLIRKNRMALFQQLTCVLPILDNLLKANVINKQEHDIKQKTQIPLQARELID 511
QY 228 SLLQROESSVRLMKVLALVFGPATCATLFIILKQYLQOE-----RLRLKQMOEEFOE 281
Db 512 TILVKGNAANIFKNCLKEIDSTLYKNLFVDMKMKYIPTEDYSGLSLEEQRLRLOEE--- 568
QY 282 HEAQLLSRAKPEDRESLSKACVCLSSFKSCVFLGCHVSCCTCYRALPEPKKPCICRQ 341
Db 569 -----RTCKVMDKEVSVVFIPOGHLVWCQEC---APSLRKCPCICRG 607

RESULT 10
PCT-US96-12860-2
; Sequence 2, Application PC/TUS9612860
; GENERAL INFORMATION:
; APPLICANT: TULARIK, INC.
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:

Db 569 -----RTCKVMDKEVSVVFIPOGHLVWCQEC---APSLRKCPCICRG 607
QY 342 AITRVIPLVNS 352
Db 608 IIKGTVRTFLS 618

RESULT 9
US-09-069-023-29
; Sequence 29, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-069-023-29

Query Match 7.9%; Score 144.5; DB 3; Length 618;
Best Local Similarity 21.9%; Pred. No. 1.3e-06;
Matches 68; Conservative 50; Mismatches 118; Indels 75; Gaps 13;

QY 82 QFVENCXGVIQRLTLOEHKVMNRTHLWDCSKIIHQRTNTPFDLPVPHEDGVDVAVRV 141
Db 343 EFVDEIQGRYPHLL--EQLLSTSDTTGEENADPPIH-----FGPGESSEDVAVM 391
QY 142 LKPL--DSVDLGL-----ETVVEKHPSTQSF---TDVIGHVISGERPKGIOE---TE 186
Db 392 NTPVKSALGEMGNRDLVKQTQVSKILTTGENTYKTVNDIVSALLNADEKREKEKEQAE 451
QY 187 EMLKVGATLTGVGEL-----VLDN-----NSVRLQPP---KQGMQYVLSQD--FD 227
Db 452 EMASDDLSLIRKNRMALFQQLTCVLPILDNLLKANVINKQEHDIKQKTQIPLQARELID 511
QY 228 SLLQROESSVRLMKVLALVFGPATCATLFIILKQYLQOE-----RLRLKQMOEEFOE 281
Db 512 TILVKGNAANIFKNCLKEIDSTLYKNLFVDMKMKYIPTEDYSGLSLEEQRLRLOEE--- 568
QY 282 HEAQLLSRAKPEDRESLSKACVCLSSFKSCVFLGCHVSCCTCYRALPEPKKPCICRQ 341
Db 569 -----RTCKVMDKEVSVVFIPOGHLVWCQEC---APSLRKCPCICRG 607

RESULT 10
PCT-US96-12860-2
; Sequence 2, Application PC/TUS9612860
; GENERAL INFORMATION:
; APPLICANT: TULARIK, INC.
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/12860
 FILING DATE: 06 AUG 1996
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: U.S. Serial Nos. 08/512,946 & 08/569,749
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Brezner, David J.
 REGISTRATION NUMBER: 24,774
 REFERENCE/DOCKET NUMBER: A-62464/DJB
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415)781-1989
 TELEFAX: (415)398-3249
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 618 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US96-12860-2

```

Query Match      7.9%; Score 144.5; DB 5; Length 618;
Best Local Similarity 21.9%; Pred. No. 1.3e-06;
Matches 68; Conservative 50; Mismatches 118; Indels 75; Gaps 13;

QY      82 QFVENCKVQRLTQEHKMWVWRTTHLWDCSKIIHQRTNTVPFDLVPHEHDGVDVAVRV 141
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db      343 EFVDEIQGRYPHLL--EQLLESTDTTGEENADPPIIH-----FGPGESSADAVM 391
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY      142 LKPL--DSVDLGL-----ETVYEKPHPSIQSF--TDVIGHYISGERPKQIQE---TE 186
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db      392 NTPVKSALBEMGNRDLVKGTQVSKILTTGENYKTVNDIVSALLNADEKREBEKEQAE 451
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY      187 EMLKVGATLTVGVEL-----VLDN-----NSVRLQPP---KQGMVYISSLQD--FD 227
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db      452 EMASDDLRLRKRMALFQQLTCVLPILDNLLKANVINKQEHDIKQKTIQIPQARELID 511
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY      228 SLLQRESSVRLMKVIALVGFATCATLFIILRKQYLQRE-----RLRLKMQEPEEQE 281
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db      512 TILVKGNAANAIKNCLEIKDSTLYKNLFDVKMKKIYPTEDVSGLSLEEQRLRLQEE--- 568
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY      282 HEAQLLSRAPREDRESLKSACVCLSSFKSVFLECGHVCSTCYEALPEPKKCPICRQ 341
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db      569 -----RTCKVMDKESVVFIPCGHLVVCQEC---APSLRKCPICRG 607
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

QY      342 AITRVIPLYNS 352
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db      608 IIKGTVRTFLS 618
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

```

```

RESULT 11
US-08-511-485-8
; Sequence 8, Application US/08511485
; Patent No. 5919912
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G.
; APPLICANT: Mackenzie, Alexander E.
; APPLICANT: Baird, Stephen
; TITLE OF INVENTION: MAMMALIAN IAP GENE FAMILY, PRIMERS,
; TITLE OF INVENTION: PROBES, AND DETECTION METHODS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA

```

```

? ZIP: 02110-2804
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/511,485
? FILING DATE: 04-AUG-1995
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: Clark, Paul T.
?
? REGISTRATION NUMBER: 30,162
? REFERENCE/DOCKET NUMBER: 07540/002001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617/542-5070
? TELEFAX: 617/542-8906
?
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 8:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 618 amino acids
? TYPE: amino acid
? STRANDEDNESS: not relevant
? TOPOLOGY: both
? MOLECULE TYPE: protein
?
? PS-08-511-485-8

```

Query Match	7.6%; Score 138.5; DB 2; Length 618;
Best Local Similarity	21.5%; Pred. No. 5.9e-06;
Matches	67; Conservative 50; Mismatches 119; Indels 75; Gaps 13;
QY	82 QFVENCKVLTQLQEHKMWNRTHLWNCCKLIHQRTNTVPFDLVPHEDGVDVAVRV 141 : : : : :
Db	343 EFVDEIQGRYPHLL--EQLLSTSDTTEENADPPIIH-----FGPGESSSEDAAVM 391 : : : : :
QY	142 LKPL--DSVDGL-----ETVYEKFHPSTQSF---TDVLGHVISGERPKGIOE---TE 186 : : : : :
Db	392 NTFPVKSALTEMGNRDLVKQTLSKITTTGENYKTVDNDIVSALLNDEDEKRREEKEKQAE 451 : : : : :
QY	187 EMLKVGATLTGVGEL-----VLDN-----NSVRLQP-----KGOMYVILSSQB--PD 227 : : : : :
Db	452 EMASDDLSLIRKNRMALFOQLTCVLPILDNLKANVINQEHDIIKKQTOIPLQARELID 511 : : : : :
QY	228 SLLQREBSSVRLWKVALVFGFATCATLFPILRKQYLQROE-----RLBLKOMQEERFQE 281 : : : : :
Db	512 TIWVKNGAANIPFNKLKSIDSTLYKNFLVDKNMKXIPTEDVSGLSLEEQLRRLQEE--- 568 : : : : :
QY	282 HEAQLLSRAPKPEDRESLKACVVVCLSSFKSCVFLEGCHVCSCCTCYRALPEPKPCICRQ 341 : : : : :
Db	569 -----RTCKVKMDKEVSVFIPCGHLVVCQEC---APSRLRKCPCICRG 607 : : : : :
QY	342 AITRVIPLYNS 352 : : : : :
Db	608 IIKGTVRTFLS 618 : : : : :

RESULT 12
US-09-212-971-8
; Sequence 8, Application US/09212971B
; Patent No. 6107041
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF TAPS AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009002
; CURRENT APPLICATION NUMBER: US/09/212,971B
; CURRENT FILING DATE: 1998-12-16

EARLIER APPLICATION NUMBER: 60/017,354
EARLIER FILING DATE: 1996-04-26
EARLIER APPLICATION NUMBER: 60/030,590
EARLIER FILING DATE: 1996-11-14
EARLIER APPLICATION NUMBER: 08/800,929
EARLIER FILING DATE: 1997-02-13
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
US-09-212-971-8

Query Match 7.6%; Score 138.5; DB 3; Length 618;
Best Local Similarity 21.5%; Pred. No. 5.9e-06;
Matches 67; Conservative 50; Mismatches 119; Indels 75; Gaps 13;
QY 82 QFVENCXGVIQRLTLOEHKQVNRTHLWDCSKIIHQRTNTVPFDLVPHEDGVDVAVRV 141
DB 343 EFVDEIQGRYPHLL--EQLLSTSDTTGEENADPPIIH-----FGPGESSEDAVMM 391
QY 142 LKPL--DSVDLGL-----ETVVEKHPSTQSF---TDVIGHVISGERPKGIOE---TE 186
DB 392 NTPVKSALMGFNRLDVKTLSKILTTGENYKTVNDIVSALLNAEDEKREBEKEKQAE 451
QY 187 EMLKVGATITGVGEL-----VLDN-----NSVRLQPP---KQGMYYLSSQD-FD 227
DB 452 EMASDDLIRKRNMAFQOQLTCLPDLNLLKANVINKEHDIIRKQTIPLQARELID 511
QY 228 SLLQROESSVRLWKVLALVFGFATCATLFFILRKQYLQOE-----RLRLKQMOBEFQ 281
DB 512 TIWKGNAANIIFKNCLKEIDSTLYKNLFVDKNMKYIPTEDVSGLSLEEQRLQEB--- 568
QY 282 HEAQLLSRAKPEDRESLSKACVCLSSFKSCVFLECGHVCSTCYRALPEPKKPCICRQ 341
DB 569 -----RTCKVMDKEVSVVFIPOGHLVVQEC---APSLAKCPCICRG 607
QY 342 AITRVIPLVNS 352
DB 608 IIRGTVRTFLS 618

RESULT 13
US-08-800-929A-8
Sequence 8, Application US/08800929A
Patent No. 6133437
GENERAL INFORMATION:
APPLICANT: Korneluk, Robert G
APPLICANT: Mackenzie, Alexander E
APPLICANT: Liston, Peter
APPLICANT: Baird, Stephen
APPLICANT: Tsang, Benjamin K
APPLICANT: Pratt, Christine
TITLE OF INVENTION: DETECTION AND MODULATION OF
TITLE OF INVENTION: IAPS AND NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
DISEASE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,929A
FILING DATE: 13-FEB-1997

CLASSIFICATION: 424
PRIOR APPLICATION DATA: 60/030,590
APPLICATION NUMBER: 60/030,590
FILING DATE: 14-NOV-1996
APPLICATION NUMBER: 60/017,354
FILING DATE: 26-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bieser-Brady, Kristina
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 07891/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 618 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-929A-8
Query Match 7.6%; Score 138.5; DB 3; Length 618;
Best Local Similarity 21.5%; Pred. No. 5.9e-06;
Matches 67; Conservative 50; Mismatches 119; Indels 75; Gaps 13;
QY 82 QFVENCXGVIQRLTLOEHKQVNRTHLWDCSKIIHQRTNTVPFDLVPHEDGVDVAVRV 141
DB 343 EFVDEIQGRYPHLL--EQLLSTSDTTGEENADPPIIH-----FGPGESSEDAVMM 391
QY 142 LKPL--DSVDLGL-----ETVVEKHPSTQSF---TDVIGHVISGERPKGIOE---TE 186
DB 392 NTPVKSALMGFNRLDVKTLSKILTTGENYKTVNDIVSALLNAEDEKREBEKEKQAE 451
QY 187 EMLKVGATITGVGEL-----VLDN-----NSVRLQPP---KQGMYYLSSQD-FD 227
DB 452 EMASDDLIRKRNMAFQOQLTCLPDLNLLKANVINKEHDIIRKQTIPLQARELID 511
QY 228 SLLQROESSVRLWKVLALVFGFATCATLFFILRKQYLQOE-----RLRLKQMOBEFQ 281
DB 512 TIWKGNAANIIFKNCLKEIDSTLYKNLFVDKNMKYIPTEDVSGLSLEEQRLQEB--- 568
QY 282 HEAQLLSRAKPEDRESLSKACVCLSSFKSCVFLECGHVCSTCYRALPEPKKPCICRQ 341
DB 569 -----RTCKVMDKEVSVVFIPOGHLVVQEC---APSLAKCPCICRG 607
QY 342 AITRVIPLVNS 352
DB 608 IIRGTVRTFLS 618

RESULT 14
US-09-617-053A-8
Sequence 8, Application US/09617053A
Patent No. 6300492
GENERAL INFORMATION:
APPLICANT: Korneluk, Robert G
APPLICANT: Mackenzie, Alexander E
APPLICANT: Liston, Peter
APPLICANT: Baird, Stephen
APPLICANT: Tsang, Benjamin K
APPLICANT: Pratt, Christine
TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
DISEASE
FILE REFERENCE: 07891/009003
CURRENT APPLICATION NUMBER: US/09/617,053A
CURRENT FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 08/800,929
PRIOR FILING DATE: 1997-02-13
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0


```
; SEQ ID NO 8
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-053A-8

Query Match
Best Local Similarity 7.6%; Score 138.5; DB 3; Length 618;
Matches 67; Conservative 50; Mismatches 119; Indels 75; Gaps 13;

QY 82 QFVENCCKGVIQRLTLOEHKMWNRTHLWNCDSKIIHQTNTVPFDLVPHEDGVDVAVRV 141
Db 343 EFVDEIQGRYPHLL--EQLLSTSDTTGEENADPPIH-----FGPCESSSEDVAVM 391

QY 142 LKPL--DSVDLGL-----ETVYKEHPHSIQSF---TDVIGHYISGERPKGIQE----TE 186
Db 392 NTPVKSALFEMGNRDLVKQTVLSKILTTGENTYKTVNDIVSALLNAEDEKREEKEKQAE 451

QY 187 EMLKVGATLTGVGEL-----VLDN-----NSVRLQPP---KQGMYYLSSOD-FD 227
Db 452 EMASDDLSLRKRNRMALFQQLTCVLPILDNLLKANVINKQEHDIHKQTKIPIQARELID 511

QY 228 SLQRQESSVRLMKVLAIVFGFATCATLFFILKQYLQROE-----RLKQMQEEOF 281
Db 512 TIWVGNAANAIFKNCLKEIDSTLYKNLFVDKMKKYIPTEDVSGLSLEQLRLQEE--- 568

QY 282 HEAQLLSRAKPEDRESLSKACVCLSFKSCVFLGCHVCSCTECYRALPEPKKCPICRQ 341
Db 569 -----RTCKVCMDEKSVVFPICGHLVVCQEC---APSLRKCPCICRG 607

QY 342 AITRVIPLVNS 352
Db 608 IIKGTVRTFLS 618

Search completed: November 6, 2004, 17:43:09
Job time : 41 secs
```

```
RESULT 15
US-09-201-936-8
; Sequence 8, Application US/09201936
; Patent No. 6541457
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G.
; APPLICANT: MacKenzie, Alexander E.
; APPLICANT: Baird, Stephen
; APPLICANT: Liston, Peter
; TITLE OF INVENTION: MAMMALIAN IAP GENE FAMILY, PRIMERS,
; FILE REFERENCE: 07891/003003
; CURRENT APPLICATION NUMBER: US/09/201,936
; EARLIER FILING DATE: 1998-12-01
; EARLIER APPLICATION NUMBER: 09/011,356
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: PCT/IB96/01022
; EARLIER FILING DATE: 1996-08-05
; EARLIER APPLICATION NUMBER: 08/576,956
; EARLIER FILING DATE: 1995-12-22
; EARLIER APPLICATION NUMBER: 08/511,485
; EARLIER FILING DATE: 1995-08-04
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-201-936-8

Query Match
Best Local Similarity 7.6%; Score 138.5; DB 4; Length 618;
Matches 67; Conservative 50; Mismatches 119; Indels 75; Gaps 13;

QY 82 QFVENCCKGVIQRLTLOEHKMWNRTHLWNCDSKIIHQTNTVPFDLVPHEDGVDVAVRV 141
Db 343 EFVDEIQGRYPHLL--EQLLSTSDTTGEENADPPIH-----FGPCESSSEDVAVM 391
```

